

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:23:06 ; Search time 11.0618 Seconds
(without alignments)
476.116 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLLIFVLATFCVFSFAQ.....DVAMGSSALGFELSFKSY 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	8.9	243	4	US-09-134-001C-2961
2	77.5	8.6	392	4	US-09-071-035-260
3	77.5	8.6	513	4	US-09-291-922-20
4	77.5	8.6	1638	4	US-09-071-035-258
5	77.5	8.6	1638	4	US-09-071-035-262
6	77.5	8.6	1638	4	US-09-071-035-266
7	74.5	8.3	222	4	US-09-134-001C-3521
8	72	8.0	332	1	US-08-118-270-53
9	72	8.0	332	5	PCT-US93-08528-53
10	71.5	7.9	300	4	US-09-134-001C-4385
11	71.5	7.9	303	4	US-09-134-001C-3009
12	71.5	7.9	655	4	US-09-134-001C-5531
13	71.5	7.9	858	4	US-09-255-829-22
14	71.5	7.9	858	4	US-09-255-829-29
15	71.5	7.9	1169	4	US-09-255-829-20
16	71	7.9	266	2	US-08-225-480-7
17	71	7.9	266	2	US-09-097-759-4
18	71	7.9	266	4	US-09-118-445-7
19	71	7.9	299	2	US-09-097-759-3
20	71	7.9	306	2	US-09-097-759-2
21	70.5	7.8	402	1	US-08-314-309A-6
22	70.5	7.8	441	1	US-08-314-309A-4
23	70.5	7.8	696	4	US-07-757-342D-4
24	70	7.8	349	4	US-09-134-001C-4004
25	70	7.8	612	3	US-09-295-186-16
26	69.5	7.7	477	4	US-09-134-001C-3487
27	69	7.6	269	1	US-08-447-554-5

28	69	7.6	269	1	US-08-468-763-17	Sequence 17, Appl
29	69	7.6	269	1	US-08-448-160-5	Sequence 5, Appl
30	69	7.6	269	1	US-08-393-996A-17	Sequence 17, Appl
31	69	7.6	611	4	US-07-757-342D-8	Sequence 8, Appl
32	69	7.6	636	4	US-07-757-342D-7	Sequence 7, Appl
33	69	7.6	674	4	US-07-757-342D-10	Sequence 10, Appl
34	69	7.6	699	4	US-07-757-342D-2	Sequence 2, Appl
35	69	7.6	1861	2	US-08-790-912-4	Sequence 4, Appl
36	68.5	7.5	494	4	US-09-134-001C-4475	Sequence 4475, Ap
37	68	7.5	462	2	US-08-898-976-2	Sequence 2, Appl
38	68	7.5	462	2	US-08-898-976-4	Sequence 4, Appl
39	68	7.5	831	2	US-08-677-734A-11	Sequence 11, Appl
40	68	7.5	831	4	US-09-097-053-11	Sequence 11, Appl
41	67.5	7.5	682	2	US-08-436-900A-4	Sequence 4, Appl
42	67.5	7.5	834	2	US-08-677-734A-9	Sequence 9, Appl
43	67.5	7.5	834	2	US-08-677-734A-10	Sequence 10, Appl
44	67.5	7.5	834	4	US-09-097-053-9	Sequence 9, Appl
45	67.5	7.5	834	4	US-09-097-053-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-134-001C-2961
; Sequence 2961, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2961
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2961

Query Match 8.9%; Score 80; DB 4; Length 243;
Best Local Similarity 30.3%; Pred. No. 0.18;
Matches 36; Conservative 16; Mismatches 53; Indels 14; Gaps 5;
QY 60 GIGSFAOGDILGGSLILGFDVAGLILAGAYLDIKALDITKKAFOWTGKGVMLAGV 119
Db 16 GIMSFSFLITAGVAAFNPCGIAL--PSYISY-LIGGETKHSHFRATFKGLGLGA 72
QY 120 VTMAVTRLTETILPFTFANSYNRKLNSLVALGFEPSFDVAMGQSSA-LGFELSFKK 177
Db 73 MT-----TGFLTIFVLAG----LLIGGLGSALTGFIFFILSLVMGILIALLLGLMFGK 121

RESULT 2
US-09-071-035-260
; Sequence 260, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

```

RESULT 4
US-09-071-035-258
; Sequence 258, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brooks
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 258:
SEQUENCE CHARACTERISTICS:
LENGTH: 1638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-258

```

```

; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 258:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-258

Query Match      8.6%; Score 77.5; DB 4; Length 1638;
Best Local Similarity 27.7%; Pred. No. 5;
Matches 43; Conservative 22; Mismatches 59; Indels 31; Gaps

QY    21  ANDSKN-----GAFGMSAGEKLLVYETSKODPIVPFLNLFLFGFIGSGFAOGDILG-- 71
Db    151  ANDLTNVITQTSGALGNLTGVDDIDLTEVRNQLELVNNIENL---GAASFTAPETLAAD 1311

QY    72  GSLLIGFDVAGIGILLIAGLDIKALDGIRKKAAFQWTGWGYMLAGVVMTAVTRLTEII 1311
Db    207  GSYISAPISDGLGLVLQAQNSNI--LQDL--NAVQALEAKGTIPS--NLVAAAANAAL 260

QY    132  LPPTFANSYNRKLNKLNSUNVALGGFEPSFDVAMGQS 166
Db    261  LP-----VKGTNNVASGALPL--LAVGGS 283

```

QY 132 LPTIFANSINKAKNSLNVALGGTFFSFVATQGS 100
 || :|||: | | :|: |
 Db 261 LP-----VKGTNVAVSGALPL--LAVGGS 283

RESULT 5

```

; ZIP: 20850
; COMPUTER TYPEABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/071,035
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 266:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-266

Query Match      8.6%; Score 77.5; DB 4; Length 1638;
Best Local Similarity 27.7%; Pred. No. 5;
Matches 43; Conservative 22; Mismatches 59; Indels 31; Gaps

QY   21 ANDSKN-----GAGCMSAGEKLLVYETSKDPIVPFLNLNLFGLFGIGSFAQGDLG-- 71
    ||| || | | | | | | | | : | : | : | : | | | | | | | | | |
Db   151 ANDLTNVITGITSGALGNLTGYDIDLTEVNRQLVNNIENL----GAASFTAPETLAAD 206
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   72 GSILGFDAVGIGILLAGAYLDIKALDGITKKAAFQTWKGVMLAGVVTMAVTRLTEII 131
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   207 GSYISAPISDGLGLVAQNYSNI--LQDL--NAAVQALEAKGTSPS--NLVAAAINAAL 260
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   132 LPPTFFANSYRKRKLNSLNVALGGFPESFDVAMQS 166
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   261 LP-----VKGTNVNAVSGALPL--LAVGGS 283
    || | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
US-09-134-001C-3521
; Sequence 3521, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO S
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3521
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3521

Query Match      8.3%; Score 74.5; DB 4; Length 222;
Best Local Similarity 25.2%; Pred. No. 0.69;
Matches 35; Conservative 17; Mismatches 50; Indels 37; Gaps

```

Db 59 FGKLSGVITIANIOTVPVIAMLMLSMGLM-----NTVFTVFLVALLPIIK 110
QY 89 GAYLDKALDKITKKAQFQWTKGVMAGVVTMAVTR---LREIILPPTTFANSYNRKLK 145
Db 111 NNTYGINEVDPNKA-----GKG-----MGTRNQVLTWIELPL----- 145
QY 146 NSLNVALGGFEPSPFDVAMG 164
Db 146 -SLSVIIGGIRIALVVAIG 163

RESULT 8

US-08-118-270-53
; Sequence 53, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2A
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-118-270-53

Query Match 8.0%; Score 72; DB 1; Length 332;
Best Local Similarity 28.7%; Pred. No. 2.4;
Matches 41; Conservative 19; Mismatches 41; Indels 42; Gaps 8;

QY 37 LLVYETSKODPIVP-FLNLFLGFGIGSPAQGDILGSLILGFDVAVGIGLILAGAYLDIK 95
Db 23 LFLVLLTSRYKLTVPREIMNL-----SFA-----DFCMLYLL-----IA 56
QY 96 ALDGITK-----KAAFWTKGKVMAGVVT-----MAVTRLEIIL-----PPTFANSYNR 142
Db 57 SVDSQTKGYNYNHAIDWQGTGCGTAGFTVLASELSVTLVITLERWHTITYAIHQ 116
QY 143 K--LKNSLNVALGGFEPSPFDVAM 163
Db 117 KLRRLHAILMLGGWLFSSLIAM 139

RESULT 9

PCT-US93-08528-53
; Sequence 53, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-53

Query Match 8.0%; Score 72; DB 5; Length 332;
Best Local Similarity 28.7%; Pred. No. 2.4;
Matches 41; Conservative 19; Mismatches 41; Indels 42; Gaps 8;

QY 37 LLVYETSKODPIVP-FLNLFLGFGIGSPAQGDILGSLILGFDVAVGIGLILAGAYLDIK 95
Db 23 LFLVLLTSRYKLTVPREIMNL-----SFA-----DFCMLYLL-----IA 56
QY 96 ALDGITK-----KAAFWTKGKVMAGVVT-----MAVTRLEIIL-----PPTFANSYNR 142
Db 57 SVDSQTKGYNYNHAIDWQGTGCGTAGFTVLASELSVTLVITLERWHTITYAIHQ 116
QY 143 K--LKNSLNVALGGFEPSPFDVAM 163
Db 117 KLRRLHAILMLGGWLFSSLIAM 139

RESULT 10

US-09-134-001C-4385
; Sequence 4385, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964

;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 4385
;; LENGTH: 300
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4385

Query Match 7.9%; Score 71.5; DB 4; Length 300;
Best Local Similarity 18.8%; Pred. No. 2.4;
Matches 51; Conservative 30; Mismatches 67; Indels 123; Gaps 11;
QY 1 MNKLIFVLATFCVFSSFAQANDSKNG-AFGMSAGEKLLVYETS----- 43
DB 2 MRKVLIFAIAGF-----LAQLVDGSLGMSGFASSSILLIYGIPAVVSAIVHSEIATT 56
QY 44 -----KODPI-----VPLLNLFLGFGIGSFAQGD----- 68
DB 57 AASGTSHWRDENVHKPTMLKLAIPGSIASAFAGAGVLFHGDYIKPFIALFLLSMGFYIL 116
QY 69 -----ILGSLILIGF-DAVGIG-----LILAGAVLDIKAL 97
DB 117 YQFLFKRAHEHHHVGNSLSFKVLPQGFVAGFLDAIGGGGWPVNTPLLSKKKIOPYIA 176
QY 98 DG-----ITKKAFAQWTWKGK-----MLAGVVTMAVTRLEIILPFTFAN 138
DB 177 IGTVSASEFFVTSSAALSFIIFLGVTQINWFVATLSLGGMVAAPISAYLVKVLPI----- 232
QY 139 SYNKKLNLSNLVALGFGFSPFDVAMGQSSAL 169
DB 233 -----NILAICVGGI-----IIFTSNAL 251

RESULT 11
US-09-134-001C-3009
; Sequence 3009, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3009
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3009

Query Match 7.9%; Score 71.5; DB 4; Length 303;
Best Local Similarity 21.0%; Pred. No. 2.4;
Matches 41; Conservative 31; Mismatches 66; Indels 57; Gaps 9;
QY 1 MNKLIF-----VLATFCVFS-----SFAQANDSKNGAFGMSA 33
DB 108 MNPLVFMVLVLLGWKVAIVFVLAIFSIFTGLVFSKMNLAEISKGVNWKDGFDFANKE 167
QY 34 GEKLLVYETSKQDP---IVPFLNLFLGFGIGSFAQGD-----LGGSLILGFDVAV 81
DB 168 GSR---FKQALNDAWAFYPLMPLVLFIGVFIAGFIYGVFPFTFTTKYASGDGIVSVFIGS 224
QY 82 GIGLILAGVLDIKALDGIYTKKAQFQWTKGVMAGVVTMAV-----TRITEIILPFTFA 137
DB 225 VIGIPM---VIRPETMLPIAEALV-----SKGMSLGTWVALLIGGASIPPEVVL---LT 273

QY 138 NSYNRKLKLSNLVAL 152
DB 274 KLFKKKFWSEFVIAI 288
RESULT 12
US-09-134-001C-5531
; Sequence 5531, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5531
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5531

Query Match 7.9%; Score 71.5; DB 4; Length 655;
Best Local Similarity 24.4%; Pred. No. 7;
Matches 42; Conservative 29; Mismatches 74; Indels 27; Gaps 7;
QY 15 FSSPAQA--NDSKNGAFGMSAGEKLLVYETSKQDPVLPFLNLFLGFGIG---SFAQGD 69
DB 349 YNAFAEQLWNIGKNSAFAL-----IIPILAG-FIARSIADKPGPAAG-L 390
QY 70 LGGSLILGFDVAVGIGLILAG---AYLDIKALDGIYTKKAQFQWTKGVMAGVVTMAVTR 126
DB 391 VGGMLAISGGSGFGIGIAGFLAGYL-TQGIKYITRKLPOAIEGLKPLIYPLLSVITG 449
QY 127 LTEIILPFTFANSYNRKLKLSNLVALGFGFSPFDVAMGQSSALGFEISFKKS 178
DB 450 LLMVYVFNPPAAWLNHLLNLGSLGSLNIMLLGLLVIGAMMAIDMGGPFNKA 501

RESULT 13
US-09-255-829-22
; Sequence 22, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:

UNITED STATES DEPARTMENT OF JUSTICE
(USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:13:37 ; Search time 29.7835 seconds
(without alignments)
7815.311 Million cell updates/sec

Title: US-09-508-487-18
Perfect score: 759
Sequence: 1 attgttaaagaattgaaat.....attctgtattgaatggggtg 759

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	44.8	5.9	6243	2	US-09-056-075-1
C 2	44.4	5.8	2032	4	US-09-241-581B-5
C 3	44.4	5.8	2032	4	US-08-265-428-5
C 4	44.4	5.8	2032	5	PCT-US95-07721-5
C 5	43	5.7	6243	2	US-09-056-075-1
C 6	41.6	5.5	5852	1	US-07-867-106-2
C 7	41.4	5.5	2547	3	US-08-262-220-7
C 8	41.4	5.5	2547	3	US-08-471-733-7
C 9	41.4	5.5	2547	3	US-08-468-878-7
C 10	41.4	5.5	2547	4	US-08-750-494-7
C 11	40.8	5.4	2251	4	US-08-991-677-11
C 12	40.8	5.4	2251	2	US-08-473-553A-1
C 13	40	5.3	998	4	US-09-122-400B-5
C 14	39.8	5.2	796	4	US-08-936-165A-103
C 15	39.2	5.2	2973	1	US-08-451-715A-7
C 16	39	5.1	237	3	US-08-680-726A-59
C 17	39	5.1	237	3	US-09-092-409-59
C 18	39	5.1	3077	3	US-09-410-028-3
C 19	39	5.1	10592	1	US-08-680-726A-51
C 20	39	5.1	10592	1	US-08-680-726A-52
C 21	39	5.1	10592	3	US-09-092-409-51
C 22	39	5.1	10592	3	US-09-092-409-52
C 23	38.8	5.1	1648	3	US-09-165-241-2
C 24	38.8	5.1	9636	1	US-08-323-170B-1
C 25	38.8	5.1	9636	4	US-08-954-441-1
C 26	38.6	5.1	1850	3	US-08-617-860B-32
C 27	38.6	5.1	4098	2	US-08-605-106-4

28	38.6	5.1	20674	4	US-09-641-638-651	Sequence 651, Appl
29	38.4	5.1	1103	4	US-09-122-400B-20	Sequence 20, Appl
30	38.4	5.1	1639	2	US-08-468-819-77	Sequence 77, Appl
C 31	38.2	5.0	1119	4	US-09-178-973B-7	Sequence 7, Appl
C 32	38.2	5.0	1119	4	US-09-419-568F-7	Sequence 7, Appl
C 33	38.2	5.0	1119	4	US-09-354-243B-7	Sequence 8, Appl
C 34	38.2	5.0	7445	4	US-09-178-973B-8	Sequence 8, Appl
C 35	38.2	5.0	7445	4	US-09-419-568F-8	Sequence 8, Appl
C 36	38.2	5.0	7445	4	US-09-354-243B-8	Sequence 8, Appl
C 37	37.8	5.0	651	4	US-09-134-001C-1749	Sequence 1749, Ap
C 38	37.8	5.0	1599	1	US-08-143-219-27	Sequence 27, Appl
C 39	37.8	5.0	19124	2	US-08-487-826B-13	Sequence 13, Appl
C 40	37.6	5.0	828	4	US-08-998-416-538	Sequence 538, App
C 41	37.6	5.0	1512	2	US-08-370-581-1	Sequence 1, Appl
C 42	37.6	5.0	1512	4	US-09-146-893-1	Sequence 1, Appl
C 43	37.6	5.0	11093	2	US-08-723-306-5	Sequence 5, Appl
C 44	37.6	5.0	11093	5	PCT-US96-10041-5	Sequence 5, Appl
C 45	37.4	4.9	2015	3	US-08-633-993A-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-09-056-075-1/c
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite
; APPLICANT: Rood, Julian
; TITLE OF INVENTION: Expression System for Clostridium
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles & Brady
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/056,075
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95238
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
; OTHER INFORMATION: plasmid RP4"
; US-09-056-075-1

Query Match 5.9%; Score 44.8; DB 2; Length 6243;
Best Local Similarity 48.7%; Pred. No. 0.2;

	Matches	153;	Conservative	0;	Mismatches	157;	Indels	4;	Gaps	1;
Qy	74	ATAATTTGAAAAAGT	TAAATTTAAATAAACCTTTTAAAACCTTTTAAATTTTCATTAAATA	133						
Db	1445	ATAATATAAAAAAATTTTTTTTCAACITTAATAATAAAATAATTTTATATTTTTTATTT	1386							
Qy	134	TGCTACCATAGTACCGACTTTTAAATAAAGGGGTTTTATGAATAAACCTTTTAAATTTTGTT	193							
Db	1385	TTTTATATTTTATAATTTTTTATTTTATTTTATATATTTTATTTTATATATTTTATTTT	1326							
Qy	194	TTGCAACCTTTTGTGTTTTTCTTAGCTTTCCTCAAGCTAATGATCTCAAANAATGGTCGG	253							
Db	1325	ATTTTATATTTTTTATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	1265							
Qy	254	TTTGGGATGAGTGCCTGGAGAAAAACTTTTGCTTTATGAACAATAGCAAGC---	AAGATCC	309						
Db	1365	TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	1206							
Qy	310	TATGTACCATTTTATTTGAACCTTTTATTTAGGGTTTGGAAATAGGCTCCTTTGCTCAAGG	369							
Db	1205	TATTTTATTTTATTTTACCCTCATTTTTCAGCTTGATTAAGGGTACTTTGTACCTGT	1146							
Qy	370	AGATATTCTGGAG	383							
Db	1145	TCTTTTTTTTGGG	1132							

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; for the Drosophila Melanogaster
; Scavenger Receptor Class CI.1
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-241-581B-5

Query Match          5.8%; Score 44.4; DB 4; Length 2032;
Best Local Similarity 46.7%; Pred. No. 0.21;
Matches 141; Conservative 0; Mismatches 161; Indels 0; Gaps

Qy 170 ATGAATAAACCTTTAATTTTGGTTTGGCAACCTTTTGGTTTCTAGCTTTGCTCAA 229
    || || || || || || || || || || || || || || || || || || ||
Db 1538 ATCAATTGTTGTTTCTTTGTTGTGTTAAACACATTAATGTAGTTGTTGCTGGTGT 1479
    || || || || || || || || || || || || || || || || || || ||
Qy 230 GCTAATCATTTCTAAAATAGTGGCTTTGGAGTGAGTGTGGAGAAAAAATTTTGGTTTAT 289
    || || || || || || || || || || || || || || || || || || ||
Db 1478 GTTGAAGTTGTAGACGTTGTAGAAGTTGTAGACATTTGGTGGTTCGAAGTCGTTGTTGTT 1419
    || || || || || || || || || || || || || || || || || || ||
Qy 290 GAAACTACGAACGAAGATCCTATTGTACCAATTTTATTTGAACCTTTTTTTTAGGGTTTGGGA 349
    || || || || || || || || || || || || || || || || || || ||
Db 1418 GCGTTCGGCGTTGTGAAGTTGTTGTCGGTTTTTTAGTGGTTGTTCGCTTTGTGACTT 1359
    || || || || || || || || || || || || || || || || || || ||
Qy 350 ATAGGCTTCCTTGCCTCAAGAGATATCTCTGGAGGTTCTCTATTATTCCTTGGATTTGATGCG 409
    || || || || || || || || || || || || || || || || || || ||
Db 1358 GTGCCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1299
    || || || || || || || || || || || || || || || || || || ||
Qy 410 GTTGGTATAGCGCTTATACCTTGTGGGGGCTTATTTGGATATCAAAAGCCCTTGATGGTATT 469
    || || || || || || || || || || || || || || || || || || ||
Db 1298 GTCGTGCTTGGTGTGTTGTTGTTGACTTCCTCGTAGTCTGTGCTGTTGTTGATGTG 1239
    || || || || || || || || || || || || || || || || || || ||
Qy 470 AC 471
    ||
Db 1238 AC 1237

RESULT 3
US-08-265-428-5/c
; Sequence 5, Application US/08265428
; Patent No. 6429289
;
; GENERAL INFORMATION:
; APPLICANT: Krieger, Monty
; TITLE OF INVENTION: Class BI Scavenger Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,428
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT6620
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2032 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdNA
; HYPOTHETICAL: NO

```

```

; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: 40..1926
; OTHER INFORMATION: /Function - "Nucleotides 40 through 1926 encode the amino acid sequence"
; OTHER INFORMATION: for the Drosophila Melanogaster
; OTHER INFORMATION: Scavenger Receptor Class C1."
;
PCT-US95-07721-5

Query Match 5.8%; Score 44.4; DB 4; Length 2032;
Best Local Similarity 46.7%; Pred. No. 0.21;
Matches 141; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 170 ATGAATAAATCTTTAAATTTTGGCAACCTTTTGTGTTTCTAGCTTTGCTCAA 229
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1538 ATCAATTATGTTTCTTTCTTTGTTAAACACATTTATTGTAGTTGTGCGTGGTGT 1479

QY 230 GCTAATGATCTTAAATGTCGCTTTGGATGAGTCTGGAGAAAACCTTTGGTTAT 289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1478 GTTGAAGTTGTAGACGTTGTAGAAGTTGTAGACTTTGTGGTGTGAAGTCGTTGTT 1419

QY 290 GAAACTAGCAAGCAAGATCTTATGTACCATTTTATTGAACCTTTTATAGGCTTTGGA 349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1418 GGCTTCGCGCTTGTGAAGTTGTGTCGGTTTATTAGTGTGTTGCTGTTGTAGTT 1359

QY 350 ATAGGCTCTTGTCTCAAGGAGATATCTTGGAGGTTCTCTTATCTTGGATTTGATCG 409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1358 GTGCCCTTGTGTTGTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1299

QY 410 GTTGGTATAGGCTTATCTTTCGGGGGCTTATTTGGATATCAAGCGCTTATGTTAT 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1298 GTCGTGTTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1239

QY 470 AC 471
    ||
Db 1238 AC 1237

RESULT 4
PCT-US95-07721-5/c
; Sequence 5, Application PC/TUS9507721
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Class BI and CI Scavenger Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07721
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT6620
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2032 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: 40..1926
; OTHER INFORMATION: /Function - "Nucleotides 40 through 1926 encode the amino acid sequence"
; OTHER INFORMATION: for the Drosophila Melanogaster
; OTHER INFORMATION: Scavenger Receptor Class C1."
;
PCT-US95-07721-5

Query Match 5.8%; Score 44.4; DB 5; Length 2032;
Best Local Similarity 46.7%; Pred. No. 0.21;
Matches 141; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 170 ATGAATAAATCTTTAAATTTTGGCAACCTTTTGTGTTTCTAGCTTTGCTCAA 229
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1538 ATCAATTATGTTTCTTTCTTTGTTAAACACATTTATTGTAGTTGTGCGTGGTGT 1479

QY 230 GCTAATGATCTTAAATGTCGCTTTGGATGAGTCTGGAGAAAACCTTTTGGTTAT 289
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1478 GTTGAAGTTGTAGACGTTGTAGAAGTTGTAGACTTTGTGGTGTGAAGTCGTTGTT 1419

QY 290 GAAACTAGCAAGCAAGATCTTATGTACCATTTTATTGAACCTTTTATAGGCTTTGGA 349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1418 GGCTTCGCGCTTGTGAAGTTGTGTCGGTTTATTAGTGTGTTGCTGTTGTAGTT 1359

QY 350 ATAGGCTCTTGTCTCAAGGAGATATCTTGGAGGTTCTCTTATCTTGGATTTGATCG 409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1358 GTGCCCTTGTGTTGTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1299

QY 410 GTTGGTATAGGCTTATCTTTCGGGGGCTTATTTGGATATCAAGCGCTTATGTTAT 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1298 GTCGTGTTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1239

QY 470 AC 471
    ||
Db 1238 AC 1237

RESULT 5
US-09-056-075-1
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite
; APPLICANT: Rood, Julian
; TITLE OF INVENTION: Expression System for Clostridium
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056.075
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95238
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3770..4013
OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
plasmid RP4"
US-09-056-075-1

Query Match 5.7%; Score 43; DB 2; Length 6243;
Best Local Similarity 49.8%; Pred. No. 0.5;
Matches 109; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 37 AAATCAAGAGCTCTATTGGGAGCGCAATTTCAAGCAATAATTTGAAAAAGTTAAATTT 96
DB 1265 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1324
QY 97 AAATACTTTAAACCTTTTAAATTCATTAATGCTACCATAGTACCAAGTTTAA 156
DB 1325 TAAAAAATATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 1384
QY 157 TAAAGGGTTTATCAATAAATCTTTTAAATTTTGTGTTGGCAACCTTTGTTTTC 216
DB 1385 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1444
QY 217 TAGCTTTGCTCAAGCTATGATCTCTAAAAATGTCGCTT 255
DB 1445 TATAATCTTTGAGAAAGCAATATAAATAAGAGCTTT 1483

RESULT 6
US-07-867-106-2
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved plasmid vectors for cellular
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2
Query Match 5.5%; Score 41.6; DB 1; Length 5852;
Best Local Similarity 53.0%; Pred. No. 1;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 72 CAATAATTTGAAAAAGTTAAATTTAAATAACTTTTAAAAACCTTTTAAATTTCAATAA 131
DB 5513 CAATGATTTTAAAAATAAAAAATCGATACATAATTTTAAAAAACCCCTTTACATTTTAT 5572
QY 132 TATGCTACCATAGTACCAAGTTTAAATAAAGGGCTTTTATGAATAAACTTTTAAATTTTG 191
DB 5573 TTATAATCCAAATTTATACATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5632
QY 192 TTTTGCAACCTTTTGTGTTTCTAGCTTTGCTCAAGCTAATGAT 239
DB 5633 TAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 5680

RESULT 7
US-08-262-220-7/c
Sequence 7, Application US/08262220
Patent No. 6054296
GENERAL INFORMATION:
APPLICANT: BERGSTROM SVEN
APPLICANT: BARBOUR ALAN G.
TITLE OF INVENTION: NEW 66 KDA ANTIGEN FROM BORRELIA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 SEVENTH STREET, N.W.
CITY: WASHINGTON
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,220
FILING DATE: 20-JUN-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: BERGSTROM-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2547 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)


```
; TYPE: DNA
; ORGANISM: Pinus taeda
US-08-991-677-11

Query Match      5.4%; Score 40.8; DB 4; Length 2251;
Best Local Similarity 48.0%; Pred. No. 1.3;
Matches 117; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 1 ATGTTAAAGAAATGAAATGATAATTTATGTCACAAATCAAGAACCTCTATTGGGAAG 60
Db 619 ACTATAAATTCAGTTGGAATGAGTGTGTTATGATTTTAAAGATGTTAAATTTATA 678
QY 61 CGAATTTCAAGCAATAATTTGAAAAAGTTAAATTTAAATAAATTTAAAAACCTTTTAA 120
Db 679 TGAATTTAAATTTATTTTGAATAACAAAAATATATATTTGGATAAAAAATTTGTTT 738
QY 121 AATTCATTAATATGCTACCATAGTACCACTTTTAAATAAGGGGTTTTTATGAATAA 180
Db 739 AAATTTAGAGTAAAAATTTCAAAATCTAAAAATATATTAACACATATTTATTTTAA 798
QY 181 TTTAATTTTGTGTCACACCTTTTGTGTTTCTTCTAGCTTTGCTCAAGCTAATGATTC 240
Db 799 TGTGGTAAATTTATCTTATATTTAAGTTAAAAATTTAGAAAAAATTAATTTAAATTA 858
QY 241 TAAA 244
Db 859 TAAA 862

RESULT 12
US-08-473-553A-1/c
; Sequence 1, Application US/08473553A
; Patent No. 5859338
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliot M.
; APPLICANT: Clark, Steven E.
; APPLICANT: Williams, Robert W.
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,
; TITLE OF INVENTION: Transformed Plants, and Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,553A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-60886/RET/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS

; LOCATION: 2434..5037
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5117..5467
US-08-473-553A-1

Query Match      5.4%; Score 40.8; DB 2; Length 5733;
Best Local Similarity 48.3%; Pred. No. 1.5;
Matches 114; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 10 AGAATGGAATGATAATTTTATGGTCAAAATCAAGAAGCTCTATTGGGAAGCAATTTCA 69
Db 1153 AAATCCAAACCGATCAAAATTAATAGCAACTGCCGATACCAAAATATATATCGTTTGTAA 1094
QY 70 AGCAATAATTTGAAAAAGTTAAATTTAAATAAATTTAAAAACCTTTTAAATTTTCATT 129
Db 1093 TCGGTGATTCAGAAAAAAGTTCAATGATATATAAATTCGTCGTAGTGTGATATTAGGAC 1034
QY 130 AATATGCTACCATAGTACCAGCTTTTAAATAAGGGGTTTTTATGAATAAATTTTAAATTT 189
Db 1033 AATCACTAGTTTGAACATAAGATACATAAACTATTTAATGTTTGAAGTTTATAGGTTT 974
QY 190 TGTTTTGGCAACCTTTTGTGTTTTTCTAGCTTTGCTCAAGCTAATGATTTCTAAAA 245
Db 973 AGTTGTTAATACTTTTGGCTGAAAAACGTGTCATTTTGTCTTAATGATGCAAAA 918

RESULT 13
US-09-122-400B-5/c
; Sequence 5, Application US/09122400B
; Patent No. 6245974
; GENERAL INFORMATION:
; APPLICANT: Michalowski, Susan
; APPLICANT: Spiker, Steven
; TITLE OF INVENTION: MATRIX ATTACHMENT REGIONS
; FILE REFERENCE: Michalowski and Spiker
; CURRENT APPLICATION NUMBER: US/09/122,400B
; CURRENT FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 60/066,118
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: DNA
; LENGTH: 998
; ORGANISM: Nicotiana tabacum
US-09-122-400B-5

Query Match      5.3%; Score 40; DB 4; Length 998;
Best Local Similarity 46.0%; Pred. No. 1.7;
Matches 172; Conservative 0; Mismatches 200; Indels 2; Gaps 1;

QY 5 TTAAGAAGATTGAAATTCATAATTTTATGGTCAAAATCAAGAAGCTCTATTGGGAAGCAA 64
Db 733 TGAATAAATATTTTCTCAAAATAATATTTTGTAAAACTGAAAAAAGCAAGAAATAT 674
QY 65 TTTCACGCAATAATTTGAAAAAGTTAAATTTAAATAAATTTTAAAAACCTTTTAAATTT 124
Db 673 TTTCTAAAGTAATTTAAAAAATCTGAAAAAATTTAAATATTTTAACTAAAAA 614
QY 125 TCATTAATATGCTACCATAGTACCAGTTTAAATAAGGGGTTTTTATGAATAAATTTTAA 184
Db 613 AGAAGAAATATTTTCTTTTAA--GTTTTTACAAAAATATGTTTTTGAATAATTTTCA 556
QY 185 ATTTTGTGTTGGCAACCTTTTGTGTTTTTCTAGCTTTGCTCAAGCTAATGATTTCTAAA 244
Db 555 GCITTTTTTAAAGCAGTTTTTTTGTAAAAATTTAGAAAAAATATTTTTCATTTTCA 496
QY 245 AATGGTGGTTTGGGATGAGTGGAGAAAAAATCTTTTGGTTTTATGAACATAGCAAGCAA 304
Db 495 GTTTTGTAGTAGTTTTTTTCCAGATTTTACAAATAAATATGCTTTAGAAAAATTTGATTT 436
QY 305 GATCCTATTGTACCATTTTATTGAACTTTTATTAGGGTTTGGAAATAGGCTCCTTGCT 364
```

Db 435 CTCTCTTTTTCAGTTTTTACAAAACATCTCTTTTAGAAAAATATTTTCAATATTTTTCT 376
QY 365 CAAGGAGATATCT 378
Db 375 AAAATAGTTTTTT 362

RESULT 14

US-08-936-165A-103/c
; Sequence 103, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimml, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-936-165A-103

Query Match 5.2%; Score 39.8; DB 4; Length 796;
Best Local Similarity 46.7%; Pred. No. 1.8;
Matches 107; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 94 TTTAAATAACTTTTAAAAACCTTTTAAATTTTCAATATGCTACCATGACAGTTT 153
Db 684 TTCATAATAATAATTAATTAATTTGATATTTCTTTNACGTTACCANAATATACATAT 625
QY 154 TAATAAGGGGTTTTTATGAATAAACTTTTAAATTTTGTGTTGGCAACCTTTGTGTTTT 213
Db 624 ACAATCAATTAATTTCTTTAAATAAATTTGTACTTATTATTATGTCATTATTTTTAT 565

QY 214 TTCTAGCTTTGCTCAAGCTAATGATCTTAAAAATGTCGTTTGGGATGAGTCTGGAGA 273
Db 564 TATTGTTTCATTTGAGNTAAACCCCTTTTGAATCAAGCTTTTAGTAGGAAAATCCATA 505
QY 274 AAAACTTTTGGTTTATGAAGACTAGCAAGCAAGTCTATTGTACCAATTT 322
Db 504 AACCCCGGTTTTTNAAGANTAATGTCATTTNAGCNAATATTTGTCAGTTT 456

RESULT 15

US-08-451-715A-7/c
; Sequence 7, Application US/08451715A
; Patent No. 5801013
; GENERAL INFORMATION:
; APPLICANT: Tao, Jianshi
; APPLICANT: Qui, Yan
; APPLICANT: Homan, Fariba
; APPLICANT: Shen, Xiaoyu
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: Helicobacter Aminoacyl-tRNA Synthetase
; TITLE OF INVENTION: Proteins, Nucleic Acids and Strains Comprising Same
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,715A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: .435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: CPI94-25
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2973 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 219..2834
US-08-451-715A-7

Query Match 5.2%; Score 39.2; DB 1; Length 2973;
Best Local Similarity 49.5%; Pred. No. 3;
Matches 101; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 19 ATTGATAATTTTATGGTCAATCAAGAAGCTCTATTGGGAAGCGAATTTTCAAGCAATAAT 78
Db 322 ATGTCTCATTCGCCATCAATTTTCAAAATATACCCCTTATGAGAGCAAAATTTCAATAATCTTT 263
QY 79 TTGAAAAAGTTTAAATTTTAAATAACTTTTAAAAACCTTTTAAATTTTCAATTAATATGCTA 138
Db 262 TTTTCTATCTCTTCGTTGGTAGTGGGGTCTCTGTTTCATTATTATTATATATCCTA 203
QY 139 CCATGATACCAAGTTTATAAAGGGGTTTTTATGAATAAACTTTTAAATTTTGTGTTTGGC 198
Db 202 AAATAGAGCGTCTTTTCAAAAAATGTTGAATTTTGGAACGCCCTTTTGTGCTTTTACGCTTTTA 143

Oy 199 AACCTTTTGTGTTTTTCTAGCTT 222
| | | | | | | | | |
Db 142 ATTGTTGGGTATTTTTTCAAAATT 119

Search completed: March 15, 2003, 07:43:26
Job time : 60.7835 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 22:48:52 ; Search time 49.6392 Seconds
(without alignments)
10731.744 Million cell updates/sec

Title: US-09-508-487-18

Perfect score: 759

Sequence: 1 attgttaagaattgaat.....attctgttgaaatgggtg 759

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51.2	6.7	416	10	US-09-960-352-4584
2	46	6.1	640681	10	US-09-790-988-1
3	45.6	6.0	2000	9	US-09-938-842A-5061
4	44.2	5.8	513509	9	US-09-754-853A-4
5	43	5.7	3979	9	US-09-989-919-52
6	42.2	5.6	419	10	US-09-960-352-11234
7	42	5.5	5520	9	US-10-001-887-43
c 8	41.8	5.5	337	10	US-09-960-352-6036
9	41.4	5.5	2846	9	US-10-008-016-1
c 10	41	5.4	640681	10	US-09-790-988-1
11	40.8	5.4	431	10	US-09-969-373-739
12	40.8	5.4	2251	10	US-09-796-256A-11
13	40.8	5.4	127197	9	US-09-754-853A-1
14	40.6	5.3	352	10	US-09-960-352-10129
15	40.6	5.3	414	10	US-09-960-352-6528
c 16	40.6	5.3	2846	9	US-10-008-016-1
17	40.6	5.3	20966	10	US-09-776-976-7
18	40.6	5.3	20966	10	US-09-758-055-7
19	40.6	5.3	20966	10	US-09-909-547-7

c 20	40.4	5.3	416	10	US-09-960-352-4584	Sequence 4584, Ap
c 21	40.2	5.3	230	10	US-09-969-373-1411	Sequence 1411, Ap
c 22	40.2	5.3	335	10	US-09-960-352-13489	Sequence 13489, A
c 23	40.2	5.3	73308	10	US-09-934-456-2276	Sequence 2276, Ap
c 24	40	5.3	988	10	US-09-867-350-1405	Sequence 1405, Ap
c 25	40	5.3	998	10	US-09-816-894-5	Sequence 5, Appli
c 26	39.8	5.2	796	10	US-09-939-980-103	Sequence 103, App
c 27	39.8	5.2	1674	9	US-09-938-842A-3489	Sequence 3489, Ap
c 28	39.8	5.2	3935	10	US-09-925-297-141	Sequence 141, App
c 29	39.4	5.2	411	10	US-09-960-352-14521	Sequence 14521, A
c 30	39.4	5.2	2000	9	US-09-938-842A-4370	Sequence 4370, Ap
c 31	39	5.1	5690	10	US-09-773-416-13	Sequence 13, Appl
c 32	39	5.1	5690	10	US-09-764-176-9	Sequence 9, Appl1
c 33	38.8	5.1	2699	8	US-08-834-666A-3	Sequence 3, Appl1
c 34	38.8	5.1	2798	8	US-08-834-666A-1	Sequence 1, Appl1
c 35	38.6	5.1	263	10	US-09-960-352-7661	Sequence 7661, Ap
c 36	38.6	5.1	371	10	US-09-969-373-1447	Sequence 1447, Ap
c 37	38.6	5.1	981	10	US-09-974-300-4908	Sequence 4908, Ap
c 38	38.6	5.1	1008	10	US-09-974-300-4909	Sequence 4909, Ap
c 39	38.6	5.1	2000	9	US-09-938-842A-4157	Sequence 4157, Ap
c 40	38.6	5.1	2000	9	US-09-938-842A-4784	Sequence 4784, Ap
c 41	38.6	5.1	2008	9	US-10-081-051-84	Sequence 84, Appl
c 42	38.6	5.1	2172	9	US-09-966-880A-15	Sequence 15, Appl
c 43	38.6	5.1	2818	9	US-09-966-880A-7	Sequence 7, Appl1
c 44	38.6	5.1	6564	9	US-09-966-880A-10	Sequence 10, Appl
c 45	38.6	5.1	11204	9	US-09-966-880A-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1

US-09-960-352-4584

; Sequence 4584, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION A

; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 4584

; LENGTH: 416

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11

US-09-960-352-4584

Query Match 6.7%; Score 51.2; DB 10; Length 416;
Best Local Similarity 50.8%; Pred. No. 0.093;
Matches 122; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 7 AAAGAATTGAATTGATTAATTTTATGGTCAATCAATCAAGAGCTCTATTGGGAGCGAAT 66

DB 164 AAAAAAAAAAAAAAAAAAGAGATTTTAAATAATATAATAATAATAATAATTCAT 223

QY 67 TCAAGCAATAATTGCAAAAAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 126

DB 224 AAAAAAAAAAATTTTAAAGAAAAAATTTTAAATTTTAAAGAAAAAATTTTAAATTAAT 283

QY 127 ATTAATATCTACCTACCTAGTACCAGTTTAAATAAGGGGTTTATGAATAAACTTTAAT 186

DB 284 TTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 343

QY 187 TTTTGTGTTGGCAACCTTTTGTGTTTCTAGCTTGTCTCAAGCTTAATGATTTCTAAAAA 246

DB 344 AATAATATGGTGAATTTTATAAATTAATAAATTTTATAAATAAATTTTATAAATAAATAA 403

RESULT 2
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 6.1%; Score 46; DB 10; Length 640681;
Best Local Similarity 49.2%; Pred. No. 4.9;
Matches 121; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 1 ATTGTTAAAGAAATGATTAATTTATGGTCAAAATCAAGAGCTCTATTGGGAAG 60
Db 35740 ATTATAAATCTATTAATAAATCAATGTTATTTTGAATAAATAATAAAAAAGAA 35799

QY 61 CGAATTTCAAGCAATAATTTGAAAAAGCTTAAATTAATAAATTTAAAAACCTTTT 120
Db 35800 CAGATTTCAATTTACAAATTCGAATAAATTTAAATATATAATATCAATCA 35859

QY 121 AATTTCATTAATGCTACCAATAGTACCAGTTTAAATAAGGGTGTATGATGAATAA 180
Db 35860 AAAAATACGGTATGAAATCGCAGTTTATTTTAAATAAATGTTATTTTAAAA 35919

QY 181 TTTAATTTTCTTGGCAACCTTTGTTTCTTCTAGCTTTGCTCAAGCTAATGATC 240
Db 35920 TAACAGTATGATTTCTACATGATTTCAATTTTCAATTTTCAATTTTAAAT 35979

QY 241 TAAAAA 246
Db 35980 ATAAAA 35985

RESULT 3
US-09-938-842A-5061
; Sequence 5061, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 5061
; LENGTH: 2000
; TYPE: DNA

; ORGANISM: Arabidopsis thaliana
US-09-938-842A-5061

Query Match 6.0%; Score 45.6; DB 9; Length 2000;
Best Local Similarity 49.0%; Pred. No. 1.7; Indels 2; Gaps 1;
Matches 150; Conservative 0; Mismatches 154; Indels 2; Gaps 1;

QY 93 ATTTAAATAACTTTAAAAACCTTTTAAATTTCAATAATATGCTACCATAGTACCAGTT 152
Db 1312 ACTTTAATATCATTAATCAATATAAAAACTTCTCTCTAGTTTGTCTTTTCTTAAT 1371

QY 153 TTAATAAGGGGTTTATGATAAATTTTAAATTTTGGCAACCTTTTGTGTTT 212
Db 1372 TT--TACGCTTTTCTAAAAATTTGATCTATAAATTTTGTGTTTATTAATCTTT 1429

QY 213 TTTCTAGCTTTGCTCAAGCTAATGATCTAAAAAATGGGCGTTGGGATGAGTGGGAG 272
Db 1430 TTAATAATGACTAAAAAATTTATAAATTTACGAATGTTTATAAGTTCTTTTAATGAATGAT 1489

QY 273 AAAAATTTTGGTTTATGAACACTAGCAAGCAAGCTCTATTCTACCATTTTATTGAACC 332
Db 1490 AACACATGTTTATTTTGGTTCACGAATTTTTCATGTTTAAATGGTAGCAACG 1549

QY 333 TTTTATAGGTTTGGATAGCTCTTTGCTCAAGGAGATATTCTTGGAGGTTCTCTTA 392
Db 1550 TTTAGTTCATGTTTAAATGGCCACATATGAACCTTGATGTTTAAAAATGCCCCACAATTTAA 1609

QY 393 TTTCTG 398
Db 1610 AACTTG 1615

RESULT 4
US-09-754-853A-4
; Sequence 4, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Haug, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 4
; LENGTH: 513509
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111805)..(113968),(114684)..(115204)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(513509)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 318013_region_A3
US-09-754-853A-4

Query Match 5.8%; Score 44.2; DB 9; Length 513509;
Best Local Similarity 51.2%; Pred. No. 11;
Matches 103; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 18 AATTGATAATTTATGTCAAATCAAGAGCTCTATTGGGAAGCAATTTCAAGCAATAA 77
Db 108907 AATTGATGTTTATATAAAAAATTTAGATCAAAATTCATCCAAATCATTTCTTAAAAATAA 108966

QY 78 TTTGAAAAAGTTAAATTTAAATAAATTTAAAAACCTTTTAAATTTCAATTAATGCT 137

RESULT 8

US-09-960-352-6036/C
; Sequence 6036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6036
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3058-015-Q1-K1-G5
US-09-960-352-6036

Query Match 5.5%; Score 41.8; DB 10; Length 337;
Best Local Similarity 51.3%; Pred. No. 6.9;
Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 35 TCAATCAAGAGCTCTATTGGGAGCGGAATTCAGCAATAATTTGAAAAAGTTAAAT 94
DB 337 TAAAAAATAAATAAAGAAACATTAATCAACATATATACTCAATTAATATAT 278
QY 95 TTAATAAATCTTTAAACCTTTTAAAAATTCATTAAATGCTACCATAGTACCAGTTT 154
DB 277 ATATTTAACAATATAAATAAAGTAAATGATATATAAATAAATGATATTAATAT 218
QY 155 ATAAAGGGTTTTTATGAATAAATCTTTAAATTTGTTGGCAACCTTTTGTGTTTTT 214
DB 217 TAAATATTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 158
QY 215 TCTAGCTTT 223
DB 157 TTTTITTTT 149

RESULT 9

US-10-008-016-1
; Sequence 1, Application US/10008016
; Patent No. US20020173630A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: 33217, A NOVEL HUMAN AMP-BINDING ENZYME
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-110001
; CURRENT APPLICATION NUMBER: US/10/008.016
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,669
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/260,166
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2846
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (245)...(2299)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2846)
; OTHER INFORMATION: n = A,T,C or G

US-10-008-016-1

Query Match 5.5%; Score 41.4; DB 9; Length 2846;
Best Local Similarity 24.8%; Pred. No. 13;
Matches 30; Conservative 55; Mismatches 36; Indels 0; Gaps 0;
QY 18 AATTGATAATTTATGGTCAAAATCAAGAGCTCTATTGGGAGCGGAATTTCAAGCAATAA 77
DB 2651 RAWYKKWKPKTTWKMYMAAAWMTTMAAAMWMMKTMARRWKAANAAMWAAATWWWWMT 2710
QY 78 TTGTGAAAAAGTTAAATTTAAATACTTTTAAAAACCTTTTAAATTTTCATTAAATATGCT 137
DB 2711 TKTWSWRTWKWARATTTTCKMWARATTTCTTWWARATKCTYWRWRTGYTTTAAATAAAT 2770
QY 138 A 138
DB 2771 A 2771

RESULT 10

US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790.988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 5.4%; Score 41; DB 10; Length 640681;
Best Local Similarity 45.1%; Pred. No. 48;
Matches 152; Conservative 0; Mismatches 185; Indels 0; Gaps 0;
QY 38 AATCAAGAGCTCTATTGGGAGCGGAATTTCAAGCAATAATTTGAAAAAGTTAAATTTA 97
DB 624202 AATAATAGCCTTTTCTCCCAATTTTGAAGAGATTTATTTAAATATTTATCTCCATG 624143
QY 98 AATACTTTAAAAACCTTTTAAAAATTTTCAATTAATATGCTACCATAGTACCAGTTTAAAT 157
DB 624142 ACTAACAATAGCACTAATTTGATATTTAAGCTCGATATAAGATAGTATTTTTTTTG 624083
QY 158 AAAGGGTTTTTATGAATAAATCTTTTAAATTTTGGTGGCAACCTTTTGTGTTTTTCT 217
DB 624082 AACGTATTTAGTTTCTTCAAGCTTAAAGGACTTTTCTGGTAGCAATTTCTCAATGTTT 624023
QY 218 AGCTTTGCTCAAGCTAATCAATCTTAAATGCTGCGTTGGGATGAGTCTGGAGAAAAA 277
DB 624022 ATGTACATATTTGGGATTCGTTATTAATAGATGATTTTATTAATTTTAAATAAAA 623963
QY 278 CTTTGGTTTTATGAAACTAGCAAGATCTTATTTGACCATTTTATTTATGACCTTTT 337
DB 623962 ATATTTTTTTTATTTTAAAGACACAAAAATATCTTTTTTTGATGCTTATATATATTT 623903
QY 338 TTAGGTTTGAATAGGCTCTCTTTTCTCAAGAGATA 374
DB 623902 TTTGGATTAAATATATATTTATTTTACAAAAAATTA 623866
RESULT 11
US-09-969-373-739

Db	Seq ID	Seq	Length	Score	DB	Length	Query Match	Best Local Similarity	48.38	Pred. No. 38	Mismatches	114	Conservative	0	Gaps	0
Db	619	ACTATAAATTTGAGTTGGAATGAGTGTGTTTATGGGATTTTTTAAGATGTTTAAATTTATATA	678													
Qy	61	CGAATTTCAAGCAATAAATTTGAAAAAGCTTAAATTTAAATAACTTTTAAAAACCTTTTTTTA	120													
Db	679	TGTAATTAATAATTTATTTTCAATTAACAAAAATATATAATTCGATATAAAAAATGTTTGTGTT	738													
Qy	121	AATTTCAATTAATATGCTTACCATAGTACAGATTTTAATAAAGGGGTTTTTATGAATAAAT	180													
Db	739	AAATTTAGAGTAAATAATTTTCAAAATCTAAAAATAATTAACACACTATATATTTTAAAAAAT	798													
Qy	181	TTTAAATTTTGTGTTTGGCAACCTTTTGTGTTTTTTCTAGCTTTCAGCTCAAGCTAATGATTC	240													
Db	799	TGTTGGTAAATTTTATCTTATATTAAGCTTAAATTTAGAAAAATTAATTTAAATTA	858													
Qy	241	TAAA 244														
Db	859	TAAA 862														
RESULT 13																
US-09-754-853A-1																
; Sequence 1, Application US/09754853A																
; Publication No. US20030005491A1																
; GENERAL INFORMATION:																
; APPLICANT: Hauge, Brian M.																
; APPLICANT: Parnell, Laurence D.																
; APPLICANT: Parsons, Jeremy D.																
; APPLICANT: Wang, Ming Li																
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With																
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance																
; FILE REFERENCE: 38-10(15810)B																
; CURRENT APPLICATION NUMBER: US/09/754,853A																
; CURRENT FILING DATE: 2001-01-05																
; PRIOR APPLICATION NUMBER: US 60/174,880																
; PRIOR FILING DATE: 2000-01-07																
; NUMBER OF SEQ ID NOS: 1119																
; SEQ ID NO 1																
; LENGTH:																

```

Qy 269 GGAGAAAACTTTTGGTTATGAACACTAGCAACAAGATCCTATTGTACACATTTTATTG 328
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 332 TTATATTGTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 391
Qy 329 AACCTTTTTTTTAGGGTTTGGAAAT 351
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 392 TATATTTTTTTTTTTTTTTTTTTAT 414

Search completed: March 15, 2003, 14:13:38
Job time : 528.639 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:19:56 ; Search time 31.1742 Seconds
(without alignments)
765.116 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLLIFVLATFCVFSFAQ.....DVAMQSSALGFLSFKKSY 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	902	100.0	179	20 AAY19808	B. burgdorferi ant
2	902	100.0	179	20 AAY04278	Borrelia burgdorfe
3	799	88.6	158	20 AAY19809	B. burgdorferi ant
4	769.5	85.3	178	20 AAY04279	Borrelia afzelii A
5	737	81.7	177	20 AAY04280	Borrelia garinii I
6	155.5	17.2	95	20 AAY04282	p11 protein sequen
7	128	14.2	25	20 AAY04281	Borrelia burgdorfe
8	84.5	9.4	378	22 AAB88515	Haemophilus influe
9	84.5	9.4	378	23 AAU91446	Haemophilus influe
10	84	9.3	355	23 ABB93914	Herbicidally activ

11	84	9.3	1456	22	ABG07111	Novel human diagno
12	84	9.3	1456	22	ABG24514	Novel human diagno
13	83.5	9.3	384	22	AAG89852	C glutamicum prote
14	83.5	9.3	384	22	AAB78951	C. glutamicum SRT
15	82	9.1	369	22	AAB78952	C. glutamicum SRT
16	82	9.1	553	20	AAY01651	A protein with cat
17	82	9.1	553	21	AAB20579	Mouse OCTN1 amino
18	81.5	9.0	572	21	AAY81595	Streptococcus pneu
19	81	9.0	325	23	ABP25872	Streptococcus poly
20	81	9.0	504	23	ABP27031	Streptococcus poly
21	80.5	8.9	111	23	ABP09270	Human ORFX protein
22	80.5	8.9	650	22	AAU37748	Streptococcus pneu
23	80	8.9	243	23	ABP38116	Staphylococcus epi
24	80	8.9	487	22	AAG91121	C glutamicum prote
25	80	8.9	487	22	AAB78989	C. glutamicum SRT
26	79.5	8.8	319	22	AAU38331	Salmonella typhi c
27	79.5	8.8	345	22	AAB76781	Corynebacterium gl
28	79.5	8.8	417	22	AAG93096	C glutamicum prote
29	79	8.8	324	22	AAB76751	Corynebacterium gl
30	79	8.8	346	22	AAG92057	C glutamicum prote
31	79	8.8	346	22	AAB76750	Corynebacterium gl
32	78.5	8.7	513	22	AAG93106	C glutamicum prote
33	77.5	8.6	392	20	AAU00139	Enterococcus faeca
34	77.5	8.6	392	23	ABP43358	E faecalis EF068 a
35	77.5	8.6	502	20	AAW99600	E stationis multi-
36	77.5	8.6	513	23	AAU97210	Corn Beta vulgaris
37	77.5	8.6	1638	20	AAU00138	Enterococcus faeca
38	77.5	8.6	1638	20	AAU00140	Enterococcus faeca
39	77.5	8.6	1638	20	AAU00142	Enterococcus faeca
40	77.5	8.6	1638	23	ABP43357	E faecalis EF068 p
41	77.5	8.6	1638	23	ABP43359	E faecalis EF069 p
42	77.5	8.6	1638	23	ABP43361	E faecalis EF070 p
43	77	8.5	251	22	AAU04855	Micromonospora eve
44	76.5	8.5	1755	21	AAG53006	Arabidopsis thalia
45	76.5	8.5	1793	21	AAG53005	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAY19808
ID AAY19808 standard; Protein; 179 AA.
XX AAY19808;
XX
XX 19-JUL-1999 (first entry)
XX
XX B. burgdorferi antigenic protein, f933.aa.
XX Antigenic protein; vaccine; Lyme disease; infection; detection.
XX Borrelia burgdorferi.
XX
XX WO9859071-A1.
XX 30-DEC-1998.
XX
XX 18-JUN-1998; 98WO-US12718.
XX
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (MEDI-) MEDIMUNE INC.
XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX WPI; 1999-189980/16.
XX N-PSDB; AAX61505.

other spirochaetes; and/or (ii) hybridizes under stringent conditions to specific nucleic acid sequences but not to genomic DNA from most other spirochaetes. (I), (II) and transformed cells are useful in CC vaccines to protect against infection by *B. burgdorferi* sensu lato. CC (I), (II) and antibodies raised against (II) are used to diagnose such CC infections, in standard immunoassays or amplification/hybridization CC tests. (I) are also used to produce recombinant (II). The 13 kD protein CC is involved in virulence and is highly conserved within *B. burgdorferi* CC sensu lato, but is absent from other *Borrelia* species (e.g. those CC responsible for relapsing fever or avian borreliosis). It should provide CC a more specific and sensitive antibody response, and diagnosis, CC compared with use of whole bacteria as antigen.

Seq	sequence	179 AA;	Query Match	100.0%;	Score 902;	DB 20;	Length 179;
			Best Local Similarity	100.0%;	Pred. No. 5.3e-95;		
			Matches 179;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MNKLIIFLVATFCVFSFAQAQNDKNGAFGMSAGEKLLVYETSKQDPVFPFLNLFGLFGF 60					
Db	1	MNKLIIFLVATFCVFSFAQAQNDKNGAFGMSAGEKLLVYETSKQDPVFPFLNLFGLFGF 60					
Qy	61	IGSFAQGDILGGSLLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTKGVMAGV 120					
Db	61	IGSFAQGDILGGSLLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTKGVMAGV 120					
Qy	121	TMAVTRITEIILPTTFANSYNRKLIKNSLUNVALGFGFSPFDVAMGOSSALGPELSFKKSY 179					
Db	121	TMAVTRITEIILPTTFANSYNRKLIKNSLUNVALGFGFSPFDVAMGOSSALGPELSFKKSY 179					

RESULT 3	
AAV19809	
ID	AAV19809 standard; Protein; 158 AA.
XX	
XX	AAV19809;
XX	
DT	19-JUL-1999 (first entry)
XX	
DE	B. burgdorferi antigenic protein, t933.aa.
XX	
KW	Antigenic protein; vaccine; Lyme disease; infection; detection
XX	
OS	Borrelia burgdorferi.
XX	
PN	WO9859071-A1.
XX	
PD	30-DEC-1998.
XX	
PF	18-JUN-1998; 98WO-US12718.
XX	
PR	03-SEP-1997; 97US-0057483.
PR	20-JUN-1997; 97US-0050359.
PR	22-JUL-1997; 97US-0053344.
PR	22-JUL-1997; 97US-0053377.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(MEDI-) MEDIMMUNE INC.
XX	
PI	Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX	
DR	WPI; 1999-189980/16.
DR	N-PSDB; AAX61506.
XX	
PT	New isolated Borrelia burgdorferi nucleic acids - used to develop
PT	products for the diagnosis, prevention and treatment of disease
PT	caused by Borrelia, particularly Lyme disease
XX	
PS	Claim 12; Page 73; 275pp; English.
XX	
CC	This sequence represents a Borrelia burgdorferi (Bb) protein of
CC	invention, which is suitable for use in a vaccine. The Bb poly

The present sequence represents a *Borrelia garinii* IP90 protein. The present invention describes an isolated nucleic acid (I) that: (i) encodes a polypeptide fragment (II) immunologically reactive with rabbit polyclonal antibody raised against a 13 kDa polypeptide of *Borrelia burgdorferi* B313, but not reactive with most proteins from other spirochaetes; and/or (ii) hybridizes under stringent conditions to specific nucleic acid sequences but not to genomic DNA from most other spirochaetes. (I), (II) and transformed cells are useful in vaccines to protect against infection by *B. burgdorferi* sensu lato. (I), (II) and antibodies raised against (II) are used to diagnose such infections, in standard immunoassays or amplification/hybridization tests. (I) are also used to produce recombinant (II). The 13 kD protein is involved in virulence and is highly conserved within *B. burgdorferi* sensu lato, but is absent from other *Borrelia* species (e.g. those

encodes a polypeptide fragment (II) immunologically reactive with rabbit polyclonal antibody raised against a 13 kDa polypeptide of *Borrelia burgdorferi* B313, but not reactive with most proteins from other spirochaetes; and/or (ii) hybridizes under stringent conditions to specific nucleic acid sequences but not to genomic DNA from most other spirochaetes. (I), (II) and transformed cells are useful in vaccines to protect against infection by *B. burgdorferi* sensu lato. (I), (II) and antibodies raised against (II) are used to diagnose such infections, in standard immunoassays or amplification/hybridization tests. (I) are also used to produce recombinant (II). The 13 kD protein is involved in virulence and is highly conserved within *B. burgdorferi* sensu lato, but is absent from other *Borrelia* species (e.g. those responsible for relapsing fever or *avain borreliosis*). It should provide a more specific and sensitive antibody response, and diagnosis,


```
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 TSKQDPVFPFLNLFGLFGIGSFAQ 66
    |||||
Db 1 TSKQDPVFPFLNLFGLFGIGSFAQ 25

RESULT 8
AAB88515
ID AAB88515 standard; Protein: 378 AA.
XX AC AAB88515;
XX DT 04-JUN-2001 (first entry)
XX DE Haemophilus influenzae essential bacterial protein SEQ ID NO:48.
XX KW Haemophilus influenzae; essential bacterial gene; identification;
XX KW otitis media; meningitis; upper respiratory tract infection;
XX KW infection; antimicrobial.
XX PS Haemophilus influenzae.
XX PN WO200111033-A2.
XX PD 15-FEB-2001.
XX PF 03-AUG-2000; 2000WO-US21176.
XX PR 04-AUG-1999; 99US-0368382.
XX PA (ABBO ) ABBOTT LAB.
XX PI Chovan LE, Hessler PE, Reich KA;
XX DR WPI: 2001-147511/15.
XX DR N-PSDB; AAF94368.
XX PT Essential bacterial genes from Haemophilus influenzae and methods for
XX PT identifying 'essential' genes that may be potential therapeutic targets
XX PS Claim 9; Page 88-89; 185pp; English.
XX CC AAF94345 to AAF94409 represent essential bacterial genes from
XX CC Haemophilus influenzae, which encode the proteins given in AAB88492 to
XX CC AAB88556. The present invention also describes methods for identifying
XX CC essential bacterial genes (i.e. those essential to the survival of a
XX CC bacterium) using a transposition system. The methods are used to
XX CC identify essential genes from bacteria, especially H. influenzae (which
XX CC causes otitis media, meningitis and upper respiratory tract infections)
XX CC which may be used as targets for potential antimicrobial agents.
XX CC AAF94410 to AAF94416 represent PCR primers used in the exemplification
XX CC of the present invention.
XX SQ Sequence 378 AA;

Query Match 9.4%; Score 84.5; DB 22; Length 378;
Best Local Similarity 28.3%; Pred. No. 0.51;
Matches 34; Conservative 19; Mismatches 48; Indels 19; Gaps 5;

Qy 14 VFSSFAQANDSKNGAFSGAGEKLLVYETSKODPIVFP-----LLNLFGLFGIGSFAQ 67
    |||||
Db 73 VKSFGLSGDGKTAIIEMAAASGLHVPPEKRNPLTTSTYGTGELIKALDLGVESFIIG 132
    :|||
Qy 68 DILGSSLILGFDVAVGIGLILAGAYIDIKALDGITKKAQFOWTWGKVMLAGVVTNAVTRL 127
    :|||
Db 133 --IGGS---ATNDGGVGMQLA---LGMQCLDSQDKPIGF-----GGAELANIVKIDVQOL 179

RESULT 9
AAB93914
XX ID AAB93914 standard; Protein: 355 AA.
XX AC AAB93914;
```

```
ID AAU91446 standard; Protein: 378 AA.
XX AC AAU91446;
XX DT 18-JUN-2002 (first entry)
XX DE Haemophilus influenzae essential gene #24.
XX KW Essential bacterial gene; antifungal agent; antibacterial agent;
XX KW antiparasitic agent; insecticidal agent; microbial infection;
XX KW mucous membrane infection; otitis media; sinusitis; bronchitis;
XX KW alveolitis; conjunctivitis; pneumonia; meningitis; epiglottitis;
XX KW cellulitis; septic arthritis.
XX OS Haemophilus influenzae.
XX PN WO200218601-A2.
XX PD 07-MAR-2002.
XX PF 22-AUG-2001; 2001WO-US26245.
XX PR 25-AUG-2000; 2000US-0649145.
XX PA (ABBO ) ABBOTT LAB.
XX PI Chovan LE, Hessler PE, Reich KA;
XX DR WPI: 2002-304258/34.
XX DR N-PSDB; ABK64932.
XX PT Essential bacterial genes in Haemophilus influenzae necessary for
XX PT bacterium's growth and survival, useful for screening inhibitors of
XX PT polypeptides and developing therapeutic agents e.g. antimicrobial
XX PS Claim 9; Page 89; 185pp; English.
XX CC The invention describes an essential bacterial gene (I) comprising a
XX CC purified polynucleotide isolated from Haemophilus influenzae where (I)
XX CC is essential to H. influenzae survival. The encoded polypeptide (II) is
XX CC useful for screening substances that function to inhibit essential H.
XX CC influenzae polypeptides by contacting (II) with the desired substances
XX CC and measuring the response by a screen from specific, enzyme, general,
XX CC affinity, phenotypic and binding screen. (I) and (II) are useful in
XX CC developing therapeutic agents such as antifungal, antibacterial and
XX CC antiparasitic agent, insecticidal agent, and preventive antimicrobial
XX CC agents which are effective in preventing microbial infection or useful
XX CC in treatment of that particular infection. (I) and (II) may also be
XX CC useful in treatment of mucous membrane infections such as otitis media,
XX CC sinusitis, bronchitis, alveolitis, conjunctivitis, pneumonia, meningitis,
XX CC epiglottitis, cellulitis and septic arthritis. This is the amino acid
XX CC sequence of an essential H. influenzae gene, described in the invention.
XX SQ Sequence 378 AA;

Query Match 9.4%; Score 84.5; DB 23; Length 378;
Best Local Similarity 28.3%; Pred. No. 0.51;
Matches 34; Conservative 19; Mismatches 48; Indels 19; Gaps 5;

Qy 14 VFSSFAQANDSKNGAFSGAGEKLLVYETSKODPIVFP-----LLNLFGLFGIGSFAQ 67
    |||||
Db 73 VKSFGLSGDGKTAIIEMAAASGLHVPPEKRNPLTTSTYGTGELIKALDLGVESFIIG 132
    :|||
Qy 68 DILGSSLILGFDVAVGIGLILAGAYIDIKALDGITKKAQFOWTWGKVMLAGVVTNAVTRL 127
    :|||
Db 133 --IGGS---ATNDGGVGMQLA---LGMQCLDSQDKPIGF-----GGAELANIVKIDVQOL 179

RESULT 10
AAB93914
XX ID AAB93914 standard; Protein: 355 AA.
XX AC AAB93914;
```

XX		31-MAY-2002	(first entry)	
XX	DT	Herbicidally active polypeptide SEQ ID NO 3125.		
XX	DE	Herbicidal; plant; agriculture; herbicide.		
XX	KW	Arabidopsis thaliana.		
XX	OS	WO200210210-A2.		
XX	PN	07-FEB-2002.		
XX	PD	28-AUG-2001; 2001WO-EF09892.		
XX	PF	28-AUG-2001; 2001WO-EF09892.		
XX	PR	(FARB) BAYER AG.		
XX	PN	Tietjen K, Weidler M;		
XX	DA	WI; 2002-269010/31.		
XX	DI	Identifying plant target proteins for herbicidally active compounds,		
XX	PT	comprising aligning and comparing nucleic acid or amino acid sequences		
XX	PT	from plant with nucleic acid or amino acid sequences from non-plant		
XX	PT	organisms -		
XX	XX	Claim 5; SEQ ID NO 3125; 261pp + Sequence Listing; English.		
XX	PS	The invention relates to identifying target proteins		
XX	CC	(ABB90790-ABB94016) for herbicidally active compounds, comprising		
XX	CC	aligning and comparing nucleic acid or amino acid sequences from plant		
XX	CC	with nucleic acid or amino acid sequences from non-plant organisms using		
XX	CC	suitable search parameters, where plant sequences having an E-value		
XX	CC	greater by a factor of 3 than the E-value of most similar non-plant		
XX	CC	sequences are selected. The polypeptides or nucleic acids encoding them		
XX	CC	are useful for identifying modulators. The identified modulators are		
XX	CC	useful as herbicides.		
XX	XX	Sequence 355 AA;		
SQ				
		Query Match 9.3%; Score 84; DB 23; Length 355;		
		Best Local Similarity 29.6%; Pred. No. 0.54;		
		Matches 48; Conservative 23; Mismatches 71; Indels 20; Gaps		9;
DY		6 IFVLATCFVSSFAQANDSKNGAFMGAGEK--LLVTETSKQDPIVPFLNLFLG-FGIIG 62		
Db		13 VLCUIVFPLLVSAEEENQCCKSGSAAEKASALKYK-----IIAFFSLIAGVFGVC 66		
QY		63 SFAQGDILGSLILGFDAVGILGLACAYLDI--KALDGITKKA-AFWTGWKGVMLAGV 119		
Db		67 LPIFGLTKTESNFFMYVKAFAGVILATGFVHILPDATESLTSGLGEPPWGDFPM-TGL 125		
QY		120 VTMAVTRLTETILPFTFANSY-NR----KLKNSLNVALGGPE 156		
Db		126 VAMAASILTMLI--ESFASGYLNRSLAKEGKTLPVSTGGEE 165		
RESULT 11				
ABG07111				
ID		ABG07111 standard; Protein; 1456 AA.		
XX	AC	ABG07111;		
XX	DT	13-FEB-2002 (first entry)		
XX	DE	Novel human diagnostic protein #7102..		
XX	XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW		food supplement; medical imaging; diagnostic; genetic disorder.		
XX	XX	Homo sapiens.		

AC ABG24514;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #24505.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS88701.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID NO 54873; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1456 AA;
Query Match 9.3%; Score 84; DB 22; Length 1456;
Best Local Similarity 21.1%; Pred. No. 3.5;
Matches 43; Conservative 32; Mismatches 75; Indels 54; Gaps 8;
OY 13 CVFSSFAQDKNKAGFCMSAGEKLLVYETSKQDPI--VPFLNLFL----- 57
DB 658 CLIESTRNSMDATNKAMLESVAENM---SVSDGVMLPLFLAMLPLVQGLAATADACT 713
OY 58 -----GFGIGSFAQDGLGGSLILGDAVGIGLILAGAYLDIKALDG 99
DB 714 FIPVTRQSDIYEVNVAAGSFGSVAAGD-----VLDMSQGVYVQLRRYVLSVSDG 767
OY 100 ITKKAQFQWT--WGKGVML-AGVVTMAVTRTEII-----LPFTFANSYRNLK--KNSL 148
DB 768 TSKTATFKMEDFEGQVPIRKGRNIYVYVNRKISVVDNCGSLLHSFTNAAGEQIVTCSL 827

OY 149 NVALGGFEPSFDVAMQSSALGFE 172
DB 828 NYNIGQIALSFSKAPDKGTEIAIE 851
RESULT 13
AAG89852
ID AAG89852 standard; Protein; 384 AA.
XX AAG89852;
AC AAG89852;
XX 26-SEP-2001 (first entry)
DT C glutamicum protein fragment SEQ ID NO: 3606.
XX
DE
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
XX EPI108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-0127688.
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX N-PSDB; AAH65071.
XX Novel polynucleotides derived from Corynebacterium, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
PS Claim 17; SEQ ID NO: 3606; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Corynebacterium, and identifying a homologue of a gene derived
CC from Corynebacterium. Corynebacterium are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 384 AA;
Query Match 9.3%; Score 83.5; DB 22; Length 384;
Best Local Similarity 24.3%; Pred. No. 0.68; Mismatches 24; Indels 91; Gaps 12;
Matches 56; Conservative 24;
OY 6 IFVLATFCVSSFAQDNDK-----NGAFGMSA----- 33
DB 9 VFVIGVTILLTAGWAHNFASVILIRBQDVSSVLVNGAFGIYALGLPLSLAGGLAD 68
OY 34 --GEKLLVYETSKQDPIVPFLNLFLGFGISFAQDGLGSLILGDAVGIGLILAGAY 91
DB 69 REGARMVLTG-----VLSALGNLSL-----LAFHD--GPSLLVGRFTVGLGV----- 110

Search completed: March 14, 2003, 14:24:22
Job time : 33.1742 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:16 ; Search time 12.4026 seconds
(without alignments)
1387.454 Million cell updates/sec

Title: US-09-508-487-19

Perfect score: 902

Sequence: 1 MNKLLIFVLATFCVFSFAQ.....DVAMGQSSALGFELSPKSKY 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	902	100.0	179	2 B70104	hypothetical prote
2	291	32.3	161	2 A70207	conserved hypothet
3	279.5	31.0	170	2 F70243	conserved hypothet
4	235	26.1	190	2 G70239	conserved hypothet
5	91	10.1	616	2 G98325	probable c4-dicarb
6	91	10.1	616	2 AG2957	hypothetical prote
7	87	9.6	449	2 S37867	hypothetical prote
8	86	9.5	214	2 S76538	hypothetical prote
9	86	9.5	307	2 T44611	cytochrome caa3 ox
10	86	9.5	429	2 A10485	c4-dicarboxylate t
11	85.5	9.5	463	2 H69822	sodium-glutamate s
12	85	9.4	146	2 S61393	hypothetical prote
13	85	9.4	146	2 F83814	Na+/H+ antiporter
14	84.5	9.4	378	2 C64142	conserved hypothet
15	84	9.3	357	2 H70346	undecaprenyl-phosp
16	83.5	9.3	484	2 T34251	hypothetical prote
17	82.5	9.1	265	2 JC4509	flagellar protein
18	81.5	9.0	174	2 T11491	NADH2 dehydrogenas
19	81.5	9.0	336	1 S75272	cytochrome d ubiqu
20	81	9.0	279	2 T35791	probable transmemb
21	81	9.0	403	2 C33958	hypothetical prote
22	81	9.0	508	2 D81325	probable phosphate
23	80.5	8.9	441	2 C96032	c4-dicarboxylate t
24	80.5	8.9	453	2 A33597	C4-dicarboxylate t
25	80.5	8.9	650	2 C95101	PTS system, fructo
26	80.5	8.9	650	2 D97969	phosphotransferase
27	80.5	8.9	666	2 A87577	oligopeptide trans
28	80.5	8.9	709	2 S75212	comE protein - Syn
29	80	8.9	264	2 S73384	hypothetical prote

30 80 8.9 385 2 G71246 hypothetical prote
31 80 8.9 538 2 A41901 metabolite export
32 80 8.9 866 2 C97662 hypothetical prote
33 80 8.9 866 2 AF2886 conserved hypothet
34 79.5 8.8 319 2 AC0956 phosphate transpor
35 79.5 8.8 493 2 S77612 amidophosphoribosy
36 79.5 8.8 692 2 H70362 glycogen phosphory
37 79 8.8 1090 2 S59077 cellulose 1,4-beta
38 78.5 8.7 375 2 AH2781 alcohol dehydrogen
39 78.5 8.7 375 2 H97560 alcohol dehydrogen
40 78.5 8.7 429 2 C72408 phosphoglucomutase
41 78.5 8.7 491 2 A70681 probable dcta prot
42 78 8.6 188 2 E71157 conserved hypothet
43 78 8.6 235 2 E89938 conserved hypothet
44 78 8.6 240 2 T38152 conserved hypothet
45 78 8.6 457 2 AH0898 Pfs system, galact

ALIGNMENTS

RESULT 1

B70104

hypothetical protein BB0034 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000

C:Accession: B70104

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; B

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: B70104

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-179 <KLE>

A:Cross-references: GB:AE001117; GB:AE000783; NID:q2687907; PIDN:AAC66426.1; PID:q22

A:Experimental source: strain B31

C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 100.0%; Score 902; DB 2; Length 179;

Best Local Similarity 100.0%; Pred. No. 3.5e-73;

Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFMSAGEKLLVYETSKODPIVPFLLNLFLGFG 60

Db 1 MNKLLIFVLATFCVFSFAQANDSKNGAFMSAGEKLLVYETSKODPIVPFLLNLFLGFG 60

QY 61 IGSAQGDILGGSLLILGFDVAVGIGLILAGAYLDIKALDGIITKKAQFQWTKGYMLAGVV 120

Db 61 IGSAQGDILGGSLLILGFDVAVGIGLILAGAYLDIKALDGIITKKAQFQWTKGYMLAGVV 120

QY 121 TMATVRLTEIILPTTFANSYNRKLNLSNALGFGFSPFDVAMGQSSALGFELSPKSKY 179

Db 121 TMATVRLTEIILPTTFANSYNRKLNLSNALGFGFSPFDVAMGQSSALGFELSPKSKY 179

RESULT 2

A70207

conserved hypothetical protein BBA01 - Lyme disease spirochete plasmid A/lp54

C:Species: Borrelia burgdorferi (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000

C:Accession: A70207

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; B

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: A70207

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-161 <KLE>
A:Cross-references: GB:AE000790; NID:g2690224; PIDN:AAC66226.1; PID:g2690225; TIGR:BBA01
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 32.3%; Score 291; DB 2; Length 161;
Best Local Similarity 39.5%; Pred. No. 7.9e-19;
Matches 75; Conservative 23; Mismatches 42; Indels 50; Gaps 8;

QY 4 LLIFVLTFCVFSFQAQDSKNGAFGMSAGEKLLVYETSKODPIVPFLLNLFGLFGIGS 63

Db 8 LLIFGL-TIQIPATKQTNRIEKGI-----ESFNKYDREKNPIGPFLLNLFGLFGIGS 60

QY 64 FAGQDILGSLILGFDVAVGIGLILAGAYLDIKALDGIITKAAAFQWTWKGCV-----114

Db 61 FVQGDYIGGSLVGFN-----LLGAIL-----WGTCIILNHRQTQL 96

QY 115 ---MLAGV-VTMAVTRLT-EIILPFTFANSYNRKLNLSNALGFGFSPDVAMGOSSAL 169

Db 97 TGYILGVGASWLTSTVSLPIPTFANRHNENLAKRLSAELAGFEPNFDGIN-----151

QY 170 GFELSFKSKY 179

Db 152 GFQLSFKSKY 161

RESULT 3

F70243

conserved hypothetical protein BBI31 - Lyme disease spirochete plasmid I/1p28-4
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: F70243

R:Praser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70243

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-170 <KLE>

A:Cross-references: GB:AE000789; NID:g2690079; PIDN:AAC66194.1; PID:g2690091; TIGR:BBI31

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 31.0%; Score 279.5; DB 2; Length 170;
Best Local Similarity 36.9%; Pred. No. 8.8e-18;
Matches 66; Conservative 29; Mismatches 69; Indels 15; Gaps 4;

QY 1 MNKLLIFVLTFCVFSFQAQDSKNGAFGMSAGEKLLVYETSKODPIVPFLLNLFGLFG 60

Db 7 MKKILTLILFSLTIQIFATQDKLEK---SVGSIEITMYKSEKATILAPFLLNLFGLTG 63

QY 61 IGSFAGQDILGSLILGFDVAVGIGLILAGAYLDIKALDGIITKAAAFQWTWKGVMAGVV 120

Db 64 IGSFVQGDYIGGAVLGSQGLGILCIAGNILG--HTDDETRA-----TTGHIITITIGV 116

QY 121 TMAVTRLTEIILPFTFANSYNRKLNLSNALGFGFSPDVAMGOSSALGFELSFKSKY 179

Db 117 TIASHIASLIIPFTFANRHNENLAKRLGIDTAGFEPNFDIGIS-----GFQLSFKSKY 170

RESULT 4

G70239

conserved hypothetical protein BBH41 - Lyme disease spirochete plasmid H/1p28-3

C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: G70239

R:Praser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vn
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: G70239

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-190 <KLE>

A:Cross-references: GB:AE000784; NID:g2690041; PIDN:AAC65989.1; PID:g2690045; TIGR:BB

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 26.1%; Score 235; DB 2; Length 190;
Best Local Similarity 38.9%; Pred. No. 9e-14; Indels 51; Gaps 4;
Matches 56; Conservative 25; Mismatches 25; Indels 51; Gaps 4;

QY 37 LLVYETSKQ-DPIVPFLLNLFGLFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIK 95

Db 58 LLQYEKEQLIDLWIPVMLNLFPLPGVGVQGDYIGGCTLGNLLGLTLTTG----VI 113

QY 96 ALDGIITKKAQFQWTWKGVMAGVVTMAVTRLTEIILPFTFANSYNRKLNLSNALGFG 155

Db 114 QMKNLKEPAS--ISSMILLSGMLTFGSSYLSIYLPVLFEDRYKYNLMRIIDELAGF 171

QY 156 EPSFDVAMGOSSALGFELSFKSKY 179

Db 172 EPNLIDIGMN-----GFQLSFKSKY 190

RESULT 5

G98325

probable c4-dicarboxylate transporter PA5169 [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: G98325

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Gold
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: G98325

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-616 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK90129.1; PID:g15160124; GSPDB:GN00170

C:Genetics:

A:Gene: AGR_L3113

A:Map position: linear chromosome

Query Match 10.1%; Score 91; DB 2; Length 616;
Best Local Similarity 22.2%; Pred. No. 2;
Matches 42; Conservative 33; Mismatches 76; Indels 38; Gaps 8;

QY 9 LATFCVFSFQAQDSKNGAFGMSAGEKLLVYE-----TSKODPIVP-----50

Db 301 LLTSVLFSG-ASGSSVANAFAFGASTFQPELVKHYPPAQAAGATIAATSVLDNVPFSIAF 359

QY 51 FLNLFGLFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIKALDGIITK-KAAFAQW- 108

Db 360 LILATATNLSVGSLLVGGFFAGGLM---AVCLGVAI---HLSVRSVDLPRATGAEWR 412

QY 109 -----TWGKGMVLAGVVTMAVTRLTEIILPFTFANSYNRKLNLSNALGFGFEPSEDV 161

Db 413 SAIAIPAFLGLGVVVVGIRIGIVITTEAA---ALAAVTLTLGLFGYRGLGVGRIFATFRQ 469

A:Cross-references: EMBL:AJ010111; NID:94584147; PIDN:CAB40606.1; PID:94584150
A:Experimental source: ATCC 10987
C:Genetics: ctab
C:Superfamily: heme O synthase

Query Match 9.5%; Score 86; DB 2; Length 307;
Best Local Similarity 25.5%; Pred. No. 2.7;
Matches 38; Conservative 27; Mismatches 58; Indels 26; Gaps 7;
QY 53 LNLFLGFGIGSFAQG-DILGSLILGFDVAGIGLILAGA-----YL--DIKALDGIK-- 102
Db 40 LVVTFGLALHFNGLSVNDLKLFFITVIGSLVMAGVCCNNYIDRDIDLMERTKTR 99
QY 103 -----KAAFQWTGKGVMLAGVVTMAVTRLEIIL-----PFTFANSYN--RKLNLSLN 149
Db 100 PVTGKYPGFALTEGLVILLGLFFELLTTPMAVLMFGAFTVIVLYSLTKKRYLYLN 159
QY 150 VALGGFEPFDMVAMQSS---ALGFELSF 175
Db 160 TVVGSISGAVPLIGWAAIDPSLGHPIAW 188

RESULT 10
A:Accession: A10485
C4-dicarboxylate transport protein dctA [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: A10485
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: A10485
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-429 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC93453.1; PID:g15981900; GSPDB:GN00175
C:Genetics: dctA
C:Superfamily: C4-dicarboxylate carrier protein

Query Match 9.5%; Score 86; DB 2; Length 429;
Best Local Similarity 25.0%; Pred. No. 3.9;
Matches 49; Conservative 30; Mismatches 63; Indels 54; Gaps 10;
QY 4 LLIF-VLATFCVFSSFAQANDSKNGAFMS-----AGEKLLVYETSKQDPIVPFLLNL 55
Db 78 LLFEIVSTLALLIGLVVNVVAPGV-GMNIDPATLDAKAVALYEAQSQGGIPEFLDI 136
QY 56 FLGFGIGSFAQGDILG---GSLILGFDVAGIG-----LILAGAYLDIKALD 98
Db 137 IPGSVVGAFAGSNILQVLLFAVLFGFALHRLGKGLIFNVIESFRSFRIFGVNMMRLA 196
QY 99 GYTKKAAQFTWCK-GVMLAGVVTMAVTRLEILPF-----TFA--NSYN- 141
Db 197 PLGAFGAMAFTIKYGV-----GSLVOLQLCYLTCILFVVLVLTGTAIKFNGFNI 249
QY 142 ----RKLNLSNLVALG 153
Db 250 FKEIRYKEELLVLVG 265

RESULT 11
H69822
sodium-glutamate symporter homolog yhcL - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: H69822
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bester, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari Nature 390, 249-256, 1997
A: Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Giech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardi, A: Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; M. Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Porter, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Seai, A: Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchly, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshi, A: Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A: Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis A: Reference number: A09580; MUID:98044033; PMID:9384377
A: Accession: H69822
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-463 <KUN>
A: Cross-references: GB:Z99108; GB:AL009126; NID:92633055; PIDN:CAB12741.1; PID:g2633 A: Experimental source: strain 168
C: Genetics:
A: Gene: yhcL
C: Superfamily: Bacillus subtilis sodium-glutamate symporter homolog yhcL

Query Match 9.5%; Score 85.5; DB 2; Length 463;
Best Local Similarity 24.3%; Pred. No. 4.6;
Matches 28; Conservative 25; Mismatches 43; Indels 19; Gaps 4;
QY 27 GAFGMSAGEK-----LLVYETSKQDPIVP-FLNLFLGFGIGSFAQGDILG-- 72
Db 334 GSFGLSIGQNCAGIYPAMLAMTPTVGQNFDFVFIITVIANVAISFGVAGVGGAAT 393
QY 73 -SLILGFDVAGIGLILAGAYLDIKALDITKKAQFQWTGKGVMLAGVVTMAVTR 126
Db 394 FAALLVLSLNNMPVALGLLISIEPLIDMGRTALNV-----SGSNTSGILTSKYK 444

RESULT 12
S61393
hypothetical protein 2 - Bacillus sp. (strain C-125)
C:Species: Bacillus sp.
A:Variety: strain C-125
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 15-Oct-1999
C:Accession: S61393
R:Hamamoto, T.; Hashimoto, M.; Hino, M.; Kitada, M.; Seto, Y.; Kudo, T.; Horikoshi, M.: Microbiol. 14, 939-946, 1994
A:Title: Characterization of a gene responsible for the Na(+)/H(+) antiporter system A:Reference number: S61392; MUID:95231300; PMID:7715455
A:Accession: S61393
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-146 <HAM>
A:Cross-references: EMBL:D31823; NID:g854654; PIDN:BAA06610.1; PID:d1007182; PID:g854 A: Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994

Query Match 9.4%; Score 85; DB 2; Length 146;
Best Local Similarity 28.8%; Pred. No. 1.6;
Matches 36; Conservative 19; Mismatches 36; Indels 34; Gaps 7;
QY 1 MNKLIFVLATFCVFSSFAQANDSKNGAFG--MSAGEKLLVYETSKQDPI---VPF---- 51
Db 16 VTRIVTFILTSFVSYLEFFAGHNDPGGFGIGLMTASAFLLMYLAFDWDMSKKALPFNFTA 75
QY 52 -----LNLFLGFGIGSFAQGD-----ILGSLI---LGFDVAGIGLILAG 89
Db 76 LIAIGLLAIFT--GVSSMLAGDPFLTYQYFRYQLPILGETETLTPALPD-LGIYLVWIG 132
QY 90 AYLDI 94
Db 133 IALTI 137

RESULT 13

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:21:36 ; Search time 7.37453 Seconds
(without alignments)
1006.744 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLLIFVLATFCVFSFAQ.....DVAMGSSALGFELSPKSKY 179

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	87	9.6	449	1 YKE6_YEAST	P36091 saccharomyc
2	86	9.5	429	1 DCTA_YERPE	Q8za28 versinia pe
3	85.5	9.5	463	1 YHCL_BACSU	P54596 bacillus su
4	84.5	9.4	378	1 GRK_HAEIN	P44507 haemophilus
5	82.5	9.1	265	1 FLIR_TREPA	P74932 treponema p
6	81.5	9.0	174	1 NU6M_RABIT	O79438 oryctolagus
7	81	9.0	461	1 PUCC_RHOCA	P23462 rhodobacter
8	80.5	8.9	441	1 DCTA_RHIME	P20672 rhizobium s
9	80.5	8.9	497	1 DCTA_RHISN	P31601 rhizobium s
10	80	8.9	264	1 YA96_MYCPN	P75596 mycoplasma
11	80	8.9	538	1 TCMA_STRGA	P39886 streptomyce
12	79.5	8.8	493	1 PURL_SYNP7	Q35038 synechococc
13	79.5	8.8	692	1 PHSG_AQUAE	O66932 aquifex ae
14	79	8.8	1090	1 GUXB_CELFI	P50899 cellulomona
15	78	8.6	188	1 Y460_PYRHO	O58215 pyrococcus
16	78	8.6	576	1 MOES_PIG	P26042 sus scrofa
17	77.5	8.6	151	1 YQ00_MYCPA	O9k537 mycobacteri
18	77	8.5	310	1 YQEW_BACSU	P54463 bacillus su
19	76.5	8.5	371	1 COXB_AGR14	Q59159 agrobacteri
20	76.5	8.5	380	1 OPS4_DROPS	P29404 drosophila
21	76.5	8.5	534	1 HUP3_CHLKE	Q39525 chlorella k
22	76	8.4	534	1 HUP1_CHLKE	P5686 chlorella k
23	75.5	8.4	356	1 YUXJ_BACSU	P40760 bacillus su
24	75.5	8.4	845	1 CC47_YEAST	P38132 saccharomyc
25	75.5	8.4	989	1 SMB2_MESAU	Q60560 mesocricetu
26	75	8.3	263	1 LPXA_CAMJE	Q9pim1 campylobact
27	75	8.3	347	1 RNHL_DROME	O9vpp5 drosophila
28	75	8.3	447	1 Y418_HAEIN	P44699 haemophilus
29	75	8.3	576	1 MOES_HUMAN	P26038 homo sapien
30	75	8.3	576	1 MOES_MOUSE	P26041 mus musculu
31	75	8.3	595	1 HMD2_SOLTU	Q41437 solanum tub
32	75	8.3	660	1 CCME_BRAJA	P45403 bradyrhizob
33	74.5	8.3	197	1 NQRE_NEIMA	O9jvq2 neisseria m

ALIGNMENTS

RESULT 1

ID	YKE6_YEAST	STANDARD;	PRT;	449 AA.
AC	P36091;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Hypothetical 49.6 kDa protein in ELM1-PR12 intergenic region.			
GN	YKL046C OR YKL259.			
OS	Saccharomyces cerevisiae (Baker's Yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S288C;			
RA	MEDLINE=94205268; PubMed=8154189;			
RA	Purnelle B., Tettelin H., van Dyck L., Skala J., Goffeau A.;			
RT	"The sequence of a 17.5 kb DNA fragment on the left arm of yeast			
RT	chromosome XI identifies the protein kinase gene ELM1, the DNA			
RT	primase gene PR12, a new gene encoding a putative histone and seven			
RT	new open reading frames."			
RL	Yeast 9:1379-1384(1993).			
CC	-1- SIMILARITY: TO YEAST YMR238W.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X71621; -- NOT ANNOTATED_CDS.			
DR	EMBL; Z28046; CA81881.1; --			
DR	PIR; S37867; S37867.			
DR	SGD; S0001529; YKL046C.			
DR	InterPro; IPR005198; Glyco_hydro.76.			
DR	Pfam; PF03663; Glyco_hydro.76; 1.			
DR	Hypothetical protein: Transmembrane.			
KW	TRANSMEM 6 26 POTENTIAL.			
FT	TRANSNEM 431 447			
SQ	SEQUENCE 449 AA; 49565 MW; DECFE9CAFD9579D5 CRC64;			

Query Match 9.6%; Score 87; DB 1; Length 449;

Best Local Similarity 27.2%; Pred. No. 2; Mismatches 39; Indels 16; Gaps 4;

Matches 28; Conservative 20; YKL046C

Qy 81 VGIIGLAGAYL-----DIKALDGIITKKAFTWCKGVMLAGVWTMAVTRTEILPFTF 136

Db 218 GVNLISNGTYKYVDGVSIDNCKTVTSQWYVYQGLLAG-----SAYLYNETG 268

Qy 137 ANSRNKLKNSLNVALGGFEPS--FDVA-MGQSSALGFELSPK 176

Db 269 SDLWHTRTKFEFLNASQVFFHDGIVVEAACQGNPNSCNTDQSPK 311

Db	197	PLGAGMARTYKIGV-----GSLVQLGQLICFLYLCILFLVVLVLTAKENGFNI	249
Qy	142	----RKLKLSINVALG 153	
		I : I : I : I	
Db	250	FKFIRYKEELLIVLG 265	
RESULT 3			
YHCL	BACSU	STANDARD;	PRT; 463 AA.
AC	P54596;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DE	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Hypothetical symporter yhcL		
GN	YHCL		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1423;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=168;		
RC	MEDLINE=97124185; PubMed=8969498;		
RX	Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.,		
RT	"A 22 kb DNA sequence in the csbp-glpFKD region at 75 degrees on the		
RT	Bacillus subtilis chromosome."		
RL	Microbiology 142:3021-3026(1996).		
RP	[2]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=168;		
RC	MEDLINE=98044033; PubMed=9384377;		
RX	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,		
RA	Azevedo V., Bersier M.G., Bessieres P., Bolotin A., Borchert S.,		
RA	Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,		
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,		
RA	Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,		
RA	Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,		
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,		
RA	Fritz C., Fujita Y., Fuma S., Galizzi A., Galleron N.,		
RA	Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grand G.,		
RA	Guiseppi G., Guy B.J., Haga K., Hatech J., Harwood C.R., Henaut A.,		
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,		
RA	Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,		
RA	Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,		
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,		
RA	Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,		
RA	Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,		
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,		
RA	Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,		
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,		
RA	Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,		
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,		
RA	Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,		
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,		
RA	Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,		
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,		
RA	Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,		
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,		
RA	Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,		
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus		
RT	subtilis".		
RL	Nature 390:249-256(1997).		
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).		
CC	-I- SUBCELLULAR LOCATION: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY		
CC	(SDF, TC 2.A.23).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement)		

```
CC or send an email to license@isb-sib.ch).
CC EMBL; X96983; CAA65696.1; -.
CC EMBL; Z99108; CAB12741.1; -.
CC Subtilist; BG11590; yhcL.
CC InterPro; IPR001991; Na/dico_symp.
CC Pfam; PF00375; SDF; 1.
CC PROSITE; PS00713; NA_DICARBOXYL_SYMP_1; FALSE_NEG.
CC PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; FALSE_NEG.
CC KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
CC FT TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 34 54 POTENTIAL.
CC FT TRANSMEM 73 93 POTENTIAL.
CC FT TRANSMEM 105 125 POTENTIAL.
CC FT TRANSMEM 184 204 POTENTIAL.
CC FT TRANSMEM 225 245 POTENTIAL.
CC FT TRANSMEM 262 282 POTENTIAL.
CC FT TRANSMEM 284 304 POTENTIAL.
CC FT TRANSMEM 338 358 POTENTIAL.
CC FT TRANSMEM 369 389 POTENTIAL.
CC FT TRANSMEM 394 414 POTENTIAL.
CC SEQUENCE 463 AA; 48982 MW; FA69EEAF5EC45F89 CRC64;

Query Match 9.5%; Score 85.5; DB 1; Length 463;
Best Local Similarity 24.3%; Pred. No. 2.7;
Matches 28; Conservative 25; Mismatches 43; Indels 19; Gaps 4;

Qy 27 GAFGMSAGEK-----LLVYETSKQDPIVP-FLNLFLFGIGSFAGQDILGG-- 72
Db 334 GSFGLSGGNCAGTYPAMLAAMTAPVGVQNPFDVFITVIAVVAISSFGVAGVGGGAT 393
Qy 73 -SLIIGDVAVGIGLAGYLDIKALDGIITKKAQFQWTKGKVMLAGVVTAVTR 126
Db 394 FAALLVLSLNPVALAGLLISIEPLIDMGRTALNV-----SGSMTSGILTSKTVK 444

RESULT 4
GRK_HAEIN
ID GRK_HAEIN STANDARD; PRT; 378 AA.
AC P44507;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycerate kinase (EC 2.7.1.31).
GN GLXK OR HI0091.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.;"
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: ATP + (R)-glycerate = ADP + 3-phospho-(R)-
CC glycerate.
CC -1- SIMILARITY: BELONGS TO THE GLYCERATE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X96983; CAA65696.1; -.
CC EMBL; Z99108; CAB12741.1; -.
CC Subtilist; BG11590; yhcL.
CC InterPro; IPR001991; Na/dico_symp.
CC Pfam; PF00375; SDF; 1.
CC PROSITE; PS00713; NA_DICARBOXYL_SYMP_1; FALSE_NEG.
CC PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; FALSE_NEG.
CC KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
CC FT TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 34 54 POTENTIAL.
CC FT TRANSMEM 73 93 POTENTIAL.
CC FT TRANSMEM 105 125 POTENTIAL.
CC FT TRANSMEM 184 204 POTENTIAL.
CC FT TRANSMEM 225 245 POTENTIAL.
CC FT TRANSMEM 262 282 POTENTIAL.
CC FT TRANSMEM 284 304 POTENTIAL.
CC FT TRANSMEM 338 358 POTENTIAL.
CC FT TRANSMEM 369 389 POTENTIAL.
CC FT TRANSMEM 394 414 POTENTIAL.
CC SEQUENCE 463 AA; 48982 MW; FA69EEAF5EC45F89 CRC64;

Query Match 9.4%; Score 84.5; DB 1; Length 378;
Best Local Similarity 28.3%; Pred. No. 2.7;
Matches 34; Conservative 19; Mismatches 48; Indels 19; Gaps 5;

Qy 14 VFSSFAQANDSKNGAFMSAGEKLLVYETSKQDPIVPF-----LLNLFLFGIGSFAGQ 67
Db 73 VKSFGLSGDGKTAIEMAAASGLHLVPPPEKRNPLITTSYGTGELIKLALDLGVESFIIG 132
Qy 68 DILGGSLLIGFDVAVGIGLAGYLDIKALDGIITKKAQFQWTKGKVMLAGVVTAVTR 127
Db 133 --IGGS---ATNDGGVGLQA---LGMQLDSQDKPIGF-----GGAELANIVKIDVQOL 179

RESULT 5
FLIR_TREPA
ID FLIR_TREPA STANDARD; PRT; 265 AA.
AC P74932;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar biosynthetic protein flir.
GN FLIR OR TP0716.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OC NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=96105201; PubMed=8529894;
RA Hardham J.M., Frye J.G., Stamm L.V.;
RT "Identification and sequences of the Treponema pallidum flir, fliv,
RT flip, fliq, flir and flhb' genes.;"
RL Gene 166:57-64(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uettersack T.,
RA McDonald L., Artlach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.;"
RL Science 281:375-388(1998).
CC -1- FUNCTION: ROLE IN FLAGELLAR BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE FLIR/MOPE/SPAR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; X96983; CAA65696.1; -.
CC EMBL; Z99108; CAB12741.1; -.
CC Subtilist; BG11590; yhcL.
CC InterPro; IPR001991; Na/dico_symp.
CC Pfam; PF00375; SDF; 1.
CC PROSITE; PS00713; NA_DICARBOXYL_SYMP_1; FALSE_NEG.
CC PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; FALSE_NEG.
CC KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
CC FT TRANSMEM 3 23 POTENTIAL.
CC FT TRANSMEM 34 54 POTENTIAL.
CC FT TRANSMEM 73 93 POTENTIAL.
CC FT TRANSMEM 105 125 POTENTIAL.
CC FT TRANSMEM 184 204 POTENTIAL.
CC FT TRANSMEM 225 245 POTENTIAL.
CC FT TRANSMEM 262 282 POTENTIAL.
CC FT TRANSMEM 284 304 POTENTIAL.
CC FT TRANSMEM 338 358 POTENTIAL.
CC FT TRANSMEM 369 389 POTENTIAL.
CC FT TRANSMEM 394 414 POTENTIAL.
CC SEQUENCE 463 AA; 48982 MW; FA69EEAF5EC45F89 CRC64;
```



```
FT TRANSMEM 183 202 PROBABLE.
FT DOMAIN 203 209 PERIPLASMIC (PROBABLE).
FT TRANSMEM 210 228 PROBABLE.
FT DOMAIN 229 261 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 262 281 PROBABLE.
FT DOMAIN 282 302 PERIPLASMIC (PROBABLE).
FT TRANSMEM 303 319 PROBABLE.
FT DOMAIN 320 338 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 339 355 PROBABLE.
FT DOMAIN 356 358 PERIPLASMIC (PROBABLE).
FT TRANSMEM 359 376 PROBABLE.
FT DOMAIN 377 394 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 395 415 PROBABLE.
FT DOMAIN 416 436 PERIPLASMIC (PROBABLE).
FT TRANSMEM 437 456 PROBABLE.
FT DOMAIN 457 461 CYTOPLASMIC (PROBABLE).
SQ SEQUENCE 461 AA; 48392 MW; 8FC7A8A0C549875A CRC64;

Query Match 9.0%; Score 81; DB 1; Length 461;
Best Local Similarity 29.1%; Pred. No. 6.6;
Matches 44; Conservative 17; Mismatches 54; Indels 36; Gaps 9;

Dy 6 IFVLATCVSFSSFAQANDSKNGAFGM--SAG-----EKLIVYTSKQDPIVPFLNL 55
:| | | | | : : : | | | | | : : | | |
Db 306 LFLAGTLAGFTASRV--LGNRGARPMRWSAGCTDRVPGFVAIIMSSLSIQDGIWFLAGT 363
:| | | | | : : : | | | | | : : | | |
Qy 56 F-LGFGIGSFAGQDILGSLILGFDVAGIGILAGAYLDIKALDGIYKKAQFQWTWKG 114
| : | | | | : | : | | | : | : | |
Db 364 FAVGLGIGLGHATLTATMRTAPADRIGLAL---GANGAVQA-----TAA-----GLGV 409
||||| : : | | | : | |
Qy 115 MLAGVTVMAVTRLEITLPFTFANS-----YN 141
||||| : : | | | : | |
Db 410 ALAGVVRDGL-----VALPGTFSGVGVPYN 435

RESULT 8
DCTA_RHIME
ID DCTA_RHIME STANDARD; PRT; 441 AA.
AC P20672;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C4-dicarboxylate transport protein.
GN DCTA OR RB1523 OR SMB20611.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymB (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=90008755; PubMed=2793824;
RA Jiang J., Gu B., Albright L.M., Nixon B.T.;
RT "Conservation between coding and regulatory elements of Rhizobium meliloti and Rhizobium leguminosarum dct genes.";
RL J. Bacteriol. 171:5244-5253(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=90008793; PubMed=2551890;
RA Engelke T., Jording D., Kapp D., Puehler A.;
RT "Identification and sequence analysis of the Rhizobium meliloti dctA gene encoding the C4-dicarboxylate carrier.";
RL J. Bacteriol. 171:5551-5560(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J1c10;
RX MEDLINE=92199261; PubMed=2134335;
RA Watson R.J.;
RT "Analysis of the C4-dicarboxylate transport genes of Rhizobium meliloti: nucleotide sequence and deduced products of dctA, dctB, and dctD.";
```

```
RL Mol. Plant Microbe Interact. 3:174-181(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
RN [5]
RP TOPOLOGY.
RX MEDLINE=94049668; PubMed=82321193;
RA Jording D., Puehler A.;
RT "The membrane topology of the Rhizobium meliloti C4-dicarboxylate permease (DctA) as derived from protein fusions with Escherichia coli K12 alkaline phosphatase (PhoA) and beta-galactosidase (LacZ).";
RL Mol. Genet. 241:106-114(1993).
CC -!- FUNCTION: Responsible for the transport of dicarboxylates such as succinate, fumarate, and malate from the periplasm across the inner membrane. This transport system plays an important role in the energy supply of rhizobium-legume symbionts.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- INDUCTION: By succinate, fumarate, and malate.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY (SDF, TC 2.A.23).
CC -!- CAUTION: The topology shown here is that reported by Ref.5. It contradicts that predicted by TM prediction programs and which has been used in orthologous entries.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26531; AAA26248.1; -
DR EMBL; M26399; AAA26253.1; -
DR EMBL; M26399; AAA26252.1; ALT_INIT.
DR EMBL; J03683; AAA63508.1; -
DR EMBL; J03683; AAA63509.1; ALT_INIT.
DR EMBL; AL603647; CAC49923.1; -
DR PIR; A33597; A33597.
DR PIR; S04816; S04816.
DR InterPro; IPR001991; Na/dico_symp.
DR Pfam; PF00375; SDF; 1.
DR PROSITE; PS00713; NA_DICARBOXYL_SYMP_1; 1.
DR PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport; Plasmid; Complete proteome.
FT DOMAIN 1 30 CYTOPLASMIC.
FT TRANSMEM 31 49 1.
FT DOMAIN 50 68 PERIPLASMIC.
FT TRANSMEM 69 87 2.
FT DOMAIN 88 99 CYTOPLASMIC.
FT TRANSMEM 100 118 3.
FT DOMAIN 119 149 PERIPLASMIC.
FT TRANSMEM 150 168 4.
FT DOMAIN 169 171 CYTOPLASMIC.
FT TRANSMEM 172 190 5.
FT DOMAIN 191 209 PERIPLASMIC.
FT TRANSMEM 210 228 6.
FT DOMAIN 229 241 CYTOPLASMIC.
FT TRANSMEM 242 260 7.
FT DOMAIN 261 281 PERIPLASMIC.
FT TRANSMEM 282 300 8.
FT DOMAIN 301 320 CYTOPLASMIC.
FT TRANSMEM 321 339 9.
FT DOMAIN 340 350 PERIPLASMIC.
FT TRANSMEM 351 369 10.
```

```
FT DOMAIN 370 378 CYTOPLASMIC.
FT TRANSMEM 379 398 11.
FT DOMAIN 399 405 PERIPLASMIC.
FT TRANSMEM 406 424 12.
FT DOMAIN 425 441 CYTOPLASMIC.
SQ SEQUENCE 441 AA; 46142 MW; B926FE7E3DC8B67D CRC64;

Query Match 8.9%; Score 80.5; DB 1; Length 441;
Best Local Similarity 27.1%; Pred. No. 7;
Matches 52; Conservative 25; Mismatches 72; Indels 43; Gaps 10;

QY 3 KLLIFVLA--TFCVFSSFAQANDSKNGAFGMSA-----GKLLVY-ETSKQDPPIVPFLL 53
DB 86 KAMIFYLAFSTLALVGLVAVVQPGA-GMHIDPASLDKAVATYAEKAHEQSITGFLM 144
QY 54 NLFGLFGISFAOGDIL-----GGSL-ILGFDVAVGIGLILAGAYLDIKALDGTGK 103
DB 145 NIPTTLVGAFAGEDILQVLFISVLFGLSIAIVGKKAEPVVDLQALTLPFRVLALMK 204
QY 104 A-----AFQWTGKGVMLAGVVTMVAITLPTFFA-----NSYN----- 141
DB 205 AAPIGAFGAMFTIGK-YGIASTANLAMLIGTFLTSLFVFLVGVAVRYNGFSILSLI 263
QY 142 RKLKNSLNVALG 153
DB 264 RYIKEELLVLG 275

RESULT 9
DCTA_RHISN STANDARD; PRT; 497 AA.
AC P31601; Q53214;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C4-dicarboxylate transport protein.
GN DCTA1 OR Y4VF.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92314432; PubMed=1617199;
RA van Sooten J.C., Bhuvanavari T.V., Bardin S., Stanley J.;
RT "Two C4-dicarboxylate transport systems in Rhizobium sp. NGR234:
RT rhizobial dicarboxylate transport is essential for nitrogen fixation
RT in tropical legume symbioses.";
RL Mol. Plant Microbe Interact. 5:179-186(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96389014; PubMed=8796346;
RA Freiberg C., Perret X., Broughton W.J., Rosenthal A.;
RT "Sequencing the 500-Kb GC-rich symbiotic replicon of Rhizobium sp.
RT NGR234 using dye terminators and a thermostable 'sequenase': a
RT beginning.";
RL Genome Res. 6:590-600(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- FUNCTION: Responsible for the transport of dicarboxylates such as
CC succinate, fumarate, and malate from the periplasm across the
CC inner membrane. This transport system plays an essential role in
CC the energy supply of tropical rhizobium-legume symbionts.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -!- INDUCTION: By succinate, fumarate, and malate. Expression depends
CC on the rpon sigma factor.
CC -!- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY
```

```
CC (SDF, TC 2.A.23).
CC -!- CAUTION: There are two genes for DctA in NGR234; one on the sym
CC plasmid, the other on the chromosome.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S38912; AAB22400.1; ALT_INIT.
CC EMBL; S38912; AAB22401.1; -.
CC EMBL; Z68203; CAA92421.1; ALT_INIT.
CC EMBL; AB000101; AAB91894.1; ALT_INIT.
CC InterPro: IPR001991; Na/diCO_symp.
CC Pfam: PF00375; SDF; 1.
CC PROSITE: PS00713; NA_DICARBOXYL_SYMP_1; 1.
CC PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; 1.
CC Transprot; Sugar transport; Transmembrane; Inner membrane; Symport;
CC KW plasmid.
CC FT TRANSMEM 27 45 POTENTIAL.
CC FT TRANSMEM 60 82 POTENTIAL.
CC FT TRANSMEM 95 117 POTENTIAL.
CC FT TRANSMEM 168 185 POTENTIAL.
CC FT TRANSMEM 205 227 POTENTIAL.
CC FT TRANSMEM 237 259 POTENTIAL.
CC FT TRANSMEM 348 370 POTENTIAL.
CC FT TRANSMEM 374 393 POTENTIAL.
CC FT CONFLICT 101 101 T -> A (IN REF. 1).
CC FT CONFLICT 337 337 A -> G (IN REF. 1).
CC FT CONFLICT 437 497 PVETMSGLSSQSDTVELGQKVLFGATNSADRTLGRPGGR
CC FT DSRRIAPDHSAQVFGGSL -> AGGDHVGVVKPAE (IN
CC REF. 1).
CC SQ SEQUENCE 497 AA; 52128 MW; 49E960D8845B2824 CRC64;

Query Match 8.9%; Score 80.5; DB 1; Length 497;
Best Local Similarity 26.9%; Pred. No. 7.9;
Matches 52; Conservative 26; Mismatches 70; Indels 45; Gaps 11;

QY 3 KLLIFVL--ATFCVFSSFAQANDSKNGAFGMSA-----GKLLVY-ETSKQDPPIVPFLL 53
DB 94 KAMIFYLFTSTLALLVGLVAVVQPGA-GMHIDPASLDKAKATYAEKAHEQSITGFLM 152
QY 54 NLFGLFGISFAOGDIL-----GGSL-ILGFDVAVGIGLILAGAYLDIKALDGTGK 103
DB 153 NIPTTLVGAFAGEDILQVLFISVLFGLSIAIVGKKAEPVVDLQALTLPFRVLALMK 212
QY 104 A-----AFQWTGK-GVMLAGVVTMVAITLPTFFA-----NSYN----- 141
DB 213 AAPIGAFGAMFTIGKGV--ASIANLAMLIGTFLTSLFVFLVGVAVRYNGFSIVAL 270
QY 142 -RKLKNSLNVALG 153
DB 271 IRYIKEELLVLG 283

RESULT 10
YA96_MYCPN
ID YA96_MYCPN STANDARD; PRT; 264 AA.
AC P75596;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MPN096 (R02_orf264).
GN MPN096 OR MP058.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
```

RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Hermann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO THE C-TERMINAL SECTION OF M.PNEUMONIAE MPN308.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AEO00007; AAB95706.1; -
DR Hypothetical protein; Transmembrane; Complete proteome.
KW TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 95 115 POTENTIAL.
FT TRANSMEM 146 166 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
SQ SEQUENCE 264 AA; 28961 MW; 84F1B7716ACB9928F CRC64;
Query Match 8.9%; Score 80; DB 1; Length 264;
Best Local Similarity 29.88; Pred. No. 4.6;
Matches 31; Conservative 18; Mismatches 35; Indels 20; Gaps 5;
QY 60 GIGSFAQGDIIIG-GLIILGFDVAGI-----GLILAGAYLDIKALDGIKKAAFOFTWKGKV 114
DB 130 KGKFKSGCGLLEFADLWTSTVSGIFWFLGLVLLGGLLIQIKK-----PKRWYFTTGLAV 185
QY 115 MLAGVVTWVTRLTEIILPF-----TFANSYNRKLKNSLNLV 152
DB 186 VVIGLTTLV-----MVQPFVDLGLIAVFNRSYRIVANTILIAI 224
RESULT 11
TCMA_STRGA STANDARD; PRT; 538 AA.
AC P39886;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tetracenomyces C resistance and export protein.
GN TCMA.
OS Streptomyces glaucoscens.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1907;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / ETH 22794 / GLA.0;
RX MEDLINE=92276347; PubMed=1592819;
RA Gullfoile P.G., Hutchinson C.R.;
RT "Sequence and transcriptional analysis of the Streptomyces
RT glaucoscens tcmAR tetracenomyces C resistance and repressor gene
RT loci";
RL J. Bacteriol. 174:3651-3658(1992).
CC -1- FUNCTION: RESISTANCE TO TETRACENOMYCIN C BY AN ACTIVE
CC TETRACENOMYCIN C EFFLUX SYSTEM WHICH IS PROBABLY ENERGIZED BY
CC TRANSMEMBRANE ELECTROCHEMICAL GRADIENTS.
CC -1- PATHWAY: Polyketide antibiotic tetracenomyces C biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80674; AAA67509.1; -
DR PIR; S27687; S27687.
DR PIR; A41901; A41901.
DR InterPro; IPR004638; Efflux_EmrB.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR TIGRFAMs; TIGR00711; efflux_EmrB; 1.
KW Antibiotic resistance; Antibiotic biosynthesis; Transmembrane;
KW Transport.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 239 259 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 319 339 POTENTIAL.
FT TRANSMEM 342 362 POTENTIAL.
FT TRANSMEM 371 391 POTENTIAL.
FT TRANSMEM 413 433 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
SQ SEQUENCE 538 AA; 54846 MW; DED3F28C1F22AA56 CRC64;
Query Match 8.9%; Score 80; DB 1; Length 538;
Best Local Similarity 25.7%; Pred. No. 9.4;
Matches 35; Conservative 20; Mismatches 33; Indels 48; Gaps 8;
QY 50 PFLNLFLGFCIG-----SFAQGDIIIGSLI--LGFDA-----VGILILAGAYL--DI 94
DB 145 PGKLNMAIGWSGVYGASTAAGPIIGLLVQHVGEAVFFINVPVGLAALVAGLVILDA 204
QY 95 KA-----LDGITKKAFAQFTWTKG-----GVMLAGVY-----TMA 123
DB 205 RAERAPKSFDSGVILISGAMFCLVWGLIKAPA--WGWGLDRTLGFLAAVLAAGFTLR 262
QY 124 VTRLTEIILPFTFANS 139
DB 263 ESRATEPLMLAMFRS 278
RESULT 12
PURI_SYN77 STANDARD; PRT; 493 AA.
ID PURI_SYN77 STANDARD; PRT; 493 AA.
AC Q55038;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine
DE DE phosphoribosylpyrophosphate amidotransferase) (ATase) (GPATase).
GN PURF.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96405630; PubMed=8809759;
RA Liu Y., Tsinorenas N.F., Golden S.S., Kondo T., Johnson C.H.;
RT "Circular expression of genes involved in the purine biosynthetic
RT pathway of the cyanobacterium Synecococcus sp. strain PCC 7942.";
RL Mol. Microbiol. 20:1071-1081(1996).
CC -1- CATALYTIC ACTIVITY: 5-phospho-beta-D-ribose + diphosphate +
CC L-glutamate -> L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate
CC + H(2)O.
CC -1- PATHWAY: De novo purine biosynthesis; first step.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.

CC	CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
CC	THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
CC	HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
CC	PROPERTIES (BY SIMILARITY).
CC	-1- CATALYTIC ACTIVITY: {(1,4)-alpha-D-glucosyl}(N) + phosphate ->
CC	{(1,4)-alpha-D-glucosyl}(N-1) + alpha-D-glucose 1-phosphate.
CC	-1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; AE000704; AAC06896.1; -.
CC	HSP; P00489; 3AMV.
DR	InterPro: IPR000811; CT_35.
DR	Pfam: PF003343; phosphorylase; 1.
DR	PROSITE; PS00102; PHOSPHORYLASE; 1.
KW	Transferase; Glycosyltransferase; Carbohydrate metabolism;
KW	Glycogen metabolism; Pyridoxal phosphate; Complete proteome.
FT	BINDING 586 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT	SEQUENCE 692 AA; 81158 MW; 4DC0EA2B2E58BEA4 CRC64;
CC	-----
CC	Query Match 8.8%; Score 79.5; DB 1; Length 692;
CC	Best Local Similarity 25.9%; Pred. No. 13;
CC	Matches 43; Conservative 22; Mismatches 58; Indels 43; Gaps 8;
CC	-----
QY	40 YETSKQDPIVPFL-----LNLFLGFGTSGFAQGDILGSLILGFDVAGVIGLIAGA 90
DB	: : : : : : : : : : :
DB	84 YEDYTKRPV-FLSPYGLHHTLLIYAG-GLG-FLAGDILKSSDLGFLPGVGFMYPGQ 140
QY	91 YL-----DIKALDGIKPAFAFTWCKGVMLA-----GV--VTWAVT 125
DB	: : : : : : : : : : :
DB	141 YVQRIQVDSQWEDLDQAQKQELMPVKVLDLKGKWLKCYVYVDEKVFYGVWVENVGKT 200
QY	126 RLTEIILPFTFANSYRKLKNSLNVALGPFSPFDVAMGQSSALGF 171
DB	: : : : : : : : : :
DB	201 KLYLDTNVEENTPWNREISSRLYV-----PKDLRLRQQIVLGF 240
CC	-----
CC	RESULT 14
CC	GUXB_CELFI STANDARD; PRT; 1090 AA.
CC	ID GUXB_CELFI
CC	AC P50899;
DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	01-OCT-1996 (Rel. 34, Last annotation update)
DE	Exoglucanase B precursor (EC 3.2.1.91) (Exocellobiohydrolase B)
GN	(1,4-beta-cellobiohydrolase B) (CBP120).
DE	CBHB OR CENE.
OS	Cellulomonas fimi.
OC	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC	Actinomycetales; Micrococciaceae; Cellulomonadaceae; Cellulomonas.
CC	NCBI_TaxID=1708;
RN	[1]
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 456-461.
RC	STRAIN=ATCC 484;
RX	MEDLINE=96003898; PubMed=7575482;
RT	Shen H., Gilkes N.R., Kilburn D.G., Miller R.C. Jr., Warren R.A.J.;
RA	"Cellobiohydrolase B, a second exo-cellobiohydrolase from the
RT	cellulolytic bacterium Cellulomonas fimi.";
RL	Biochem. J. 311:67-74(1995).
RN	[2]
RP	SEQUENCE OF 54-75.
RX	MEDLINE=93209933; PubMed=8458833;
RA	Meinke A., Gilkes N.R., Kilburn D.G., Miller R.C. Jr., Warren R.A.J.;
RT	"Cellose-binding polypeptides from Cellulomonas fimi: endoglucanase
RT	D (CenD), a family A beta-1,4-glucanase.";
RL	J. Bacteriol. 175:1910-1918(1993).

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:51 ; Search time 24.47 Seconds
(without alignments)
1507.249 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLIFVLATFCVFSFAQ.....DVAMGSSALGFELSFKSKY 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	902	100.0	179	16	O51065 borrelia bu
2	769.5	85.3	178	2	Q9X309
3	737	81.7	177	2	Q9X3V0
4	291	32.3	161	16	O50896
5	279.5	31.0	170	16	O50885
6	247.5	27.4	161	2	Q9S011
7	235	26.1	190	16	O50696
8	155.5	17.2	95	2	O31325
9	94.5	10.5	427	2	Q93RE6
10	91	10.1	616	16	Q8UAV6
11	87.5	9.7	198	16	Q98KK8
12	87.5	9.7	289	16	Q9XKK6
13	87.5	9.7	442	17	Q976Y7
14	87.5	9.7	462	16	Q92MI7
15	86.5	9.6	1874	2	O33763
16	86	9.5	214	16	Q55742

17	86	9.5	274	16	O8XQB4
18	86	9.5	303	2	Q93PN3
19	86	9.5	307	2	Q9XB19
20	85	9.4	146	16	O44659
21	85	9.4	995	5	Q9N3H7
22	84	9.3	355	10	Q9FIS2
23	84	9.3	357	16	O66804
24	83.5	9.3	470	5	Q19932
25	83.5	9.3	538	17	Q9HH12
26	83	9.2	420	16	Q9CP19
27	82.5	9.1	438	17	Q96XR2
28	82.5	9.1	439	2	Q9RBQ9
29	82	9.1	153	17	O8TYN4
30	82	9.1	553	11	Q9Z306
31	82	9.1	1021	17	Q8TWM0
32	81.5	9.0	296	17	Q9HL22
33	81.5	9.0	336	16	P73160
34	81	9.0	144	2	Q9RG24
35	81	9.0	279	16	O87858
36	81	9.0	427	2	O8VS09
37	81	9.0	457	2	Q9R637
38	81	9.0	508	16	Q9PNA1
39	81	9.0	553	11	Q9RI41
40	81	9.0	695	17	Q97A18
41	80.5	8.9	455	16	Q9ZS78
42	80.5	8.9	650	16	Q97RE5
43	80.5	8.9	665	16	O9A523
44	80.5	8.9	709	16	P73100
45	80.5	8.9	1914	2	O33762

ALIGNMENTS

RESULT 1

O51065	PRELIMINARY;	PRT;	179 AA.
AC	O51065;		
DT	01-JUN-1998 (TrEMBLrel. 06, Created)		
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)		
DE	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Hypothetical protein BB0034.		
GN	BB0034 OR P13.		
OS	Borrelia burgdorferi (Lyme disease spirochete).		
OC	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.		
OX	NCBI_taxid=139;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 35210 / B31;		
EX	MEDLINE=98065943; PubMed=9403685;		
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,		
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,		
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,		
RA	Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M.,		
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,		
RA	Utterback T., Matthey L., McDonald L., Artiach P., Bowman C.,		
RA	Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,		
RA	Smith H.O., Venter J.C.;		
RT	"Genomic sequence of a Lyme disease spirochaete, Borrelia		
RL	burgdorferi."		
RN	Nature 390:580-586(1997).		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 35210 / B31;		
RA	Noppa L., Ostberg Y., Bergstrom S.;		
RT	"p13, a novel integral membrane protein of Lyme disease Borrelia."		
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF001117; AAC66426.1; -		
DR	EMBL; AF085739; AAD28360.1; -		
DR	TIGR; BB0034; -		
DR	Hypothetical protein; Complete proteome.		
SQ	SEQUENCE 179 AA; 19104 MW; FD6056E8E24E5D5D CRC64;		

```
Query Match 100.0%; Score 902; DB 16; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.8e-72;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60
DB 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60

QY 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAVV 120
DB 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAVV 120

QY 121 TMAVTRTEILPFTFANSYNRKLNALGFGFSPFDVAMQSSALGFLSFKKSY 179
DB 121 TMAVTRTEILPFTFANSYNRKLNALGFGFSPFDVAMQSSALGFLSFKKSY 179

RESULT 2
Q9X309 PRELIMINARY; PRT; 178 AA.
AC Q9X309;
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACAI;
RX MEDLINE=21189251; PubMed=11292755;
RA Noppa L., Ostberg Y., Lavrinovicha M., Bergstrom S.;
RT "p13, an integral membrane protein of Borrelia burgdorferi, is C-
terminally processed and contains surface-exposed domains.";
RL Infect. Immun. 69:3323-3334(2001).
DR EMBL: AF085740; AAD28361.1; -.
SQ SEQUENCE 178 AA; 19194 MW; C1C032EB5320C86 CRC64;

Query Match 85.3%; Score 769.5; DB 2; Length 178;
Best Local Similarity 84.9%; Pred. No. 1.5e-60;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60
DB 1 MNKLLIFVLATFCVFSFAQANDSKS-AFNLGAGEKLLAYETSKQDPIVPFLLNLFGLFG 59
QY 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAVV 120
DB 60 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDKNAPKAFKWTGKGMVLGAVV 119

QY 121 TMAVTRTEILPFTFANSYNRKLNALGFGFSPFDVAMQSSALGFLSFKKSY 179
DB 120 TMAVTRTEILPFTFANSYNRKLNALGFGFSPFDINMGQASALGFLSFKKSY 178

RESULT 3
Q9X3V0 PRELIMINARY; PRT; 177 AA.
ID Q9X3V0;
AC Q9X3V0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Membrane protein p13.
GN p13.
OS Borrelia garinii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ip90;
RX MEDLINE=21189251; PubMed=11292755;
```

```
RA Noppa L., Ostberg Y., Lavrinovicha M., Bergstrom S.;
RT "p13, an integral membrane protein of Borrelia burgdorferi, is C-
terminally processed and contains surface-exposed domains.";
RL Infect. Immun. 69:3323-3334(2001).
DR EMBL: AF085741; AAD28362.1; -.
SQ SEQUENCE 177 AA; 19308 MW; 429928CC426E1111 CRC64;

Query Match 81.7%; Score 737; DB 2; Length 177;
Best Local Similarity 81.0%; Pred. No. 1.1e-57;
Matches 145; Conservative 14; Mismatches 18; Indels 2; Gaps 2;

QY 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 60
DB 1 MNKFLIFLIVFICAFSSFAQ-DDSKS-TFNLGAGERFLVYETNKKDSLVPFLLNLFGLFG 58
QY 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAVV 120
DB 59 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKDFDNNAKKADFKWTGKGMVLGAVV 118

QY 121 TMAVTRTEILPFTFANSYNRKLNALGFGFSPFDVAMQSSALGFLSFKKSY 179
DB 119 TMAVTRTEILPFTFANNYNRKLNALGFGFSPFDINMGQASALGFLSFKKSY 177

RESULT 4
O50896 PRELIMINARY; PRT; 161 AA.
ID O50896;
AC O50896;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBA01.
GN BBA01.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Utterback T., Wattley L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL: AE000790; AAC66226.1; -.
DR TIGR: BBA01; -.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 161 AA; 17887 MW; D93BDB326FE2DA30 CRC64;
```

```
Query Match 32.3%; Score 291; DB 16; Length 161;
Best Local Similarity 39.5%; Pred. No. 3e-18;
Matches 75; Conservative 23; Mismatches 42; Indels 50; Gaps 8;

QY 4 LLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLLNLFGLFG 63
DB 8 LLIFGL-TQIFATKQTQRIEKG-----ESFNKYDREKKNIPGFLNLFGLFG 60
QY 64 FAQGDILGSLILGFDVAGIGLILAGAYLDIKALDGIITKKAQFQWTGKGMVLGAVV 114
DB 61 FVQGDYIGGSLVGFN-----LLGAIL-----WGTLNHNRETQL 96
QY 115 ---MLAGV-VTMAVTRLT-EIILPFTFANSYNRKLNALGFGFSPFDVAMQSSAL 169
DB 97 TGVILGVGASWLTSTVTLIIPFTFANNHNENLKKRLSAELAGFEPNLDGIN----- 151
```


Qy 170 GFELSFKSKY 179
Db 152 GFQLSFKSKY 161

RESULT 5

OS0885 PRELIMINARY; PRT; 170 AA.
AC OS0885;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BB131.
GN BB131.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp28-4.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
RA Uterback T., Watthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi";
RL Nature 390:580-586(1997).
DR EMBL; AE000789; AAC66194.1; -
DR TIGR; BB131; -
RW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 170 AA; 16227 MW; 0E71B633EA80FB22 CRC64;

Query Match 31.0%; Score 279.5; DB 16; Length 170;
Best Local Similarity 36.9%; Pred. No. 3.4e-17;
Matches 66; Conservative 29; Mismatches 69; Indels 15; Gaps 4;
Qy 1 MNKLIFVATFCVFSFAQAANDSKNGAFSGAGEKLLVYETSKODPIVPELLNLFGLFG 60
Db 7 MKKLTILIFSLTIQIFATQDKLEK---SVGSITIMKYKSEKATILAPLLNLFGLTG 63
Qy 61 IGSAQGDILGSLILGFDVAGIGLILAGAYLDIKALDITKKAQFQWTKGKGYMLAGV 120
Db 64 IGSPVQGDYIGGGAVLGSLGLGILCIAGNLTG---HTDDETRA-----TTGHIITTTIGV 116
Qy 121 TMATVTRLEITLPTTFANSYNRKLNLSNALVGGFPEPSFDVAMGSSALGFELSKSKY 179
Db 117 TIIASHATSLIPTTFANKHNANLKKRLGIDIAGFEPNFDIGIS-----GFQLSFKKRY 170

RESULT 6

OS05011 PRELIMINARY; PRT; 161 AA.
AC OS05011;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Conserved hypothetical protein.
GN BB006.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp56.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,
RA White O., Dodson R., Hickey E.K., Gwinn M., Peterson J., van Vugt R.,
RA Palmer N., Haft D., Rosa P., Stevenson B.;

RT "A bacterial genome in flux: The twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi.",
RL Mol. Microbiol. 0:0-0(1999).
DR EMBL; AE001584; AAF07707.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 161 AA; 17016 MW; 14D4906CC8107CD2 CRC64;

Query Match 27.4%; Score 247.5; DB 2; Length 161;
Best Local Similarity 40.9%; Pred. No. 2.2e-14;
Matches 63; Conservative 25; Mismatches 51; Indels 15; Gaps 6;

Qy 18 FAQANDSKNGAFSGAGEKLLVYETSKODPIVPELLNLFGLFGIGSFAGQDILGSLILG 77
Db 4 FAQEKLEK---GVGDIATVMKYSEKATILAPLLNLFSLGIGSFVQGDYIGGALLG 59
Qy 78 FDVAGIGLILAGAYL--DIKALDITKKAQFQWTKGKGYMLAGV--VTMAYTRLEITLIP 133
Db 60 SQVLGGILIMAG-YMTGDI---GFVTESTATVITGG---VLSGIGGLTAAASYITGIIP 112
Qy 134 FTTFANSYNRKLNLSNALVGGFPEPSFDVAMGSS 167
Db 113 KFFANRYNADKKRLGIALAGLEPNFDIGINDS 146

RESULT 7

OS05096 PRELIMINARY; PRT; 190 AA.
ID OS05096
AC OS05096;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBH41.
GN BBH41.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp28-3.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
RA Uterback T., Watthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi";
RL Nature 390:580-586(1997).
DR EMBL; AE000784; AAC65989.1; -
DR TIGR; BBH41; -
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 190 AA; 21813 MW; A80FB8BCAD4B13EB CRC64;

Query Match 26.1%; Score 235; DB 16; Length 190;
Best Local Similarity 38.9%; Pred. No. 3.3e-13;
Matches 56; Conservative 25; Mismatches 51; Indels 12; Gaps 4;

Qy 37 LLVYETSKQ-DPIVPELLNLFGLFGIGSFAGQDILGSLILGFDVAGIGLILAGAYLDIK 95
Db 58 LLQYKEGLIDILWIPVMLNLFDPFGVSGVQGDYIGGCTGLGNLGLTGTG---VI 113
Qy 96 ALDGITTKKAQFQWTKGKGYMLAGVVTMAYTRLEITLPTTFANSYNRKLNLSNALVGGF 155
Db 114 QMKNLKEPAS--ISSMILLSGMLTFCGSSYLSIYLPVLPFDYRYKLMNRIIDELAGF 171
Qy 156 EPSFDVAMGQSSALGFELSKSKY 179
Db 172 EPNLDIGMNN---GFQLSFKKSY 190

```

99 GI--TKAAAFQWTKGVMVLAVVTVMAVTRLTETIILPFTFANSYNRKLNLSNALGGFE 156
187 GYLWLKHLIAFGWD--TRASIVLDKTIIRFLSLSPDTTF-----NFLAWAGLGF- 234

QY 157 PSFDVAMGOSSALG 170
DB 235 -AYHAIMGQAGVQG 247

RESULT 10
Q8UAV6
ID Q8UAV6 PRELIMINARY; PRT; 616 AA.
AC Q8UAV6;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DE ABC transporter, membrane spanning protein.
GN ATU3261 OR AGR_L_3113.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE009255; AAL44077.1; -
DR EMBL; AE008358; AAK90129.1; -
SQ Complete proteome.
KW SEQUENCE 616 AA; 63402 MW; A6005D4347ABE060 CRC64;

Query Match 10.1%; Score 91; DB 16; Length 616;
Best Local Similarity 22.2%; Pred. No. 7.4;
Matches 42; Conservative 33; Mismatches 76; Indels 38; Gaps 8;

QY 9 LATFCVFSFAQANDSKNGAFGMSAGEKLLVYE-----TSKQDPIVP----- 50
DB 301 LUTSVLFSG-ASGSSVANAFGASTFQPELVKHGYPQAAGAIATSVLDNVIPPSIAF 359
QY 51 FLINLFLGFGIGSFAQGDILGSLILGFDVAVGIGLILAGAYLDIKALDGIK- KAAQW- 108
DB 360 LILATATNLVSGSLLVGCGFFAGGLM---AVCLGVAI---HLVSRSVDTLPRTAGRW 412
QY 109 -----TWGKGVMVLAVVTVMAVTRLTETIILPFTFANSYNRKLNLSNALGGFE 161
DB 413 SIAAIPAFGLGVVVVVGIRIGIVTTTAA---ALAAALYLLGLFGYRLGVGRIFATFRQ 469
QY 162 AMGOSSALG 170
: |:::|:

QY 1 MKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPFLNLFGRF 60
DB 1 MKKIETLILFGLTIEIFATKDTQNRKIEGIESFNK---YDREKNPIGPFLLNLFPLPG 57
QY 61 IGSFAQGDILGSLILGFDPAVG 82
DB 58 IGSFVGQDYIGGSGVLGFNLLG 79

RESULT 9
Q93RE6
ID Q93RE6 PRELIMINARY; PRT; 427 AA.
AC Q93RE6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DE Adhesin.
GN Adhesin.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RA Akeda Y., Honda T.;
RT "Vibrio parahaemolyticus adhesin gene."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047560; BAB59008.1; -
DR InterPro: IPR002528; Mate.
DR Pfam: PF01554; UPP0013; 2.
DR TIGRFAMs: TIGR00797; mate; 1.
SQ SEQUENCE 427 AA; 45909 MW; 0B608B9DC70E3C71 CRC64;

Query Match 10.5%; Score 94.5; DB 2; Length 427;
Best Local Similarity 25.8%; Pred. No. 2.4;
Matches 50; Conservative 25; Mismatches 72; Indels 47; Gaps 11;

QY 5 LIFVLATFCVFSFAQ-----ANDSKNGAFGMSAGEKLLVYETSKQDPIV----- 49
DB 73 MFAALAVCLFVFFQPIINGLTDSOE---VIELGSSYLVISSASFAVACVASMAGVLR 129
QY 50 ----PFLNLFGLFGIGS--FQGDILGSLILGFDVAVG-----TGLILAGAYLDIKALD 98
DB 130 AMHQPGLSFTFFSGIGLSNVLNWLFIQGH--LGFPALGITGAATATVISA-IEVGCLF 186
```

```
Db 470 SAGEAATIG 478
RESULT 11
Q98KK8 PRELIMINARY; PRT; 198 AA.
AC Q98KK8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein mll1430.
GN Mll1430.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002997; BAB48806.1; -.
DR InterPro; IPR001123; Lyse.
DR Pfam; PF01810; Lyse; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 198 AA; 21126 MW; 67E14425F02E7DAB CRC64;

Query Match 9.7%; Score 87.5; DB 16; Length 198;
Best Local Similarity 30.0%; Pred. No. 4;
Matches 45; Conservative 17; Mismatches 51; Indels 37; Gaps 9;

Qy 49 VPFLNLFLGFGIGSFAOGDIILG-GSLILGPDVAGIGL-ILAGAYL-----DIKALDGIT 101
Db 40 VPHMLGICSGFLVLLAVG--LGLGAVLTFALHTALKIGAGVLLYLKAWKMSRLT 97
Qy 102 KK-----AAFQWTWKGVMGLAGVVTMAVTRLTE-----IILPFTFANSY 140
Db 98 GKGETKARPMRFIDAAAFQWNPRAWMA-ITAMAVYANAEPFLSVALISTAFIVNLP 156
Qy 141 NRKLKNSLUNVALGGF--EPS-----FDVANG 164
Db 157 SVSVWAGFTALRGFLSDPVRKWFNIANG 186

RESULT 12
Q9KXK6 PRELIMINARY; PRT; 289 AA.
AC Q9KXK6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC02314 OR SC53.05.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL357591; CAB93434.1; -.
SQ SEQUENCE 289 AA; 29123 MW; 89AD68D9D805BE10 CRC64;

Query Match 9.7%; Score 87.5; DB 16; Length 289;
Best Local Similarity 32.6%; Pred. No. 6.2;
Matches 31; Conservative 12; Mismatches 37; Indels 15; Gaps 4;

Qy 73 SLILGDAVGIG---LILAGAYLDIKALDGITKKAQFQWTWKGVMGLAGVVTMAVTRLTE 129
Db 5 SLTRG-DGVVIGAAVLLLIASFLDIYSIDGAPDSADISLWGSVGLVGLAGVIGAL 63
Qy 130 IILPFTFANSYNRKLKNSLUNVA---LGGFEPSPFDV 161
Db 64 VVV-----NRAMPQPKVAGLDLGGFGIAFTV 90

RESULT 13
Q976Y7 PRELIMINARY; PRT; 442 AA.
AC Q976Y7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative transporter.
GN ST0054.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000981; BAB65009.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 442 AA; 47602 MW; C7932E5ABF20F03A CRC64;

Query Match 9.7%; Score 87.5; DB 17; Length 442;
Best Local Similarity 22.8%; Pred. No. 10;
Matches 47; Conservative 29; Mismatches 75; Indels 55; Gaps 10;

Qy 4 LLIFVLATFCVF-SSFAQANDSKNGA--FGMSAGEKL-----LVYETSKQD-- 46
Db 49 VFLATAVYATFLTSLM--RPVGLYFGNSISDKLGRKRAILYIGLIVITASLQAL 105
Qy 47 -----PIVPFLNLFLGFGIGSFAQGD-----ILGGSILILGFDVAG 82
Db 106 PTYAVVGIFAPILILVRLAEFFVGGVTAGSHHTIGPESVPERHGWGG---IGFSAAG 162
Qy 83 IGLIILAGAYLDIKALDGITKKAQFQWTWKGVM-LAGVVTMAVTRLTEILP----FTFA 137
Db 163 AAYLTASAWFYMTAL---LFGSGSYLVMGWRVYMFEGGLPLAVLGVSVLYVPESDAWQKA 219
Qy 138 NSYNRKLKNSLUNVALGGFEPSPFDVAM 163
Db 220 KDRGKTIKSPVRELFKFRFTFGIAL 245
```

```
RESULT 14
Q92MI7 PRELIMINARY; PRT; 462 AA.
AC Q92MI7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical transmembrane protein SMC00743.
GN R02629 OR SMC00743.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelie D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47208.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 462 AA; 4905 MW; 0EBEB9B4CE96F14F CRC64;

Query Match 9.7%; Score 87.5; DB 16; Length 462;
Best Local Similarity 25.9%; Pred. No. 11;
Matches 37; Conservative 20; Mismatches 53; Indels 33; Gaps 4;

QY 51 FLNLFGL-EGISFAGQDILGGS-----LILGFDVAVGIGLILAG 89
Db 142 YVPLXILGLFTLPMIALGDLGDTARAHSPVIVAMSPYLVLRPLLILGFNVIAIA---AG 198

QY 90 AYLDIKALDGITKKAFOFTWKGVMIA-----GVVTMAVTRTEILPFTFANSY 140
Db 199 LPRDATTMGAALAAAYVTLTGOFVNWRLGRHYRGPMPKIELGRWLSVSLPIFLVDGF 258

QY 141 NRKLKNSLVALGGFEPFDVAM 163
Db 259 GFLTNSDVVIVGLYLPDDVAV 281

RESULT 15
O33763 PRELIMINARY; PRT; 1874 AA.
AC O33763;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE IgA1 protease (EC 3.4.24.13).
GN IGA.
OS Streptococcus sanguis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK4;
RX MEDLINE=98084472; PubMed=9423856;
RA Poulsen K., Reinholdt J., Jespersgaard C., Boye K., Brown T.A.,
RA Hauge M., Kilian M.;
RT "A comprehensive genetic study of streptococcal immunoglobulin A1
RT proteases: evidence for recombination within and between species.";
RL Infect. Immun. 66:181-190(1998).
DR EMBL; Y13459; CAA73856.1; -.
DR MEROPS; M26.001; -.
DR InterPro; IPR001899; Gram_pos_anchor.
```

```
DR InterPro; IPR000130; Zn_MTpeptdse.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRFAMS; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMS; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Hydrolase; Protease.
SQ SEQUENCE 1874 AA; 207888 MW; FA30684926BB4532 CRC64;

Query Match 9.6%; Score 86.5; DB 2; Length 1874;
Best Local Similarity 21.8%; Pred. No. 67;
Matches 41; Conservative 30; Mismatches 64; Indels 53; Gaps 6;

QY 5 LIFVLATFCVFSSFAQAANDSKNGAFGMSAG-----EKLIVYETSKQDPIVPFLLNLFGLG 60
Db 83 LVYELANSNTGAKTLPTNTGDSNNSNTMTAGLLTIGLVFVFSKRKYQSKFLLTVLYGAG 142

QY 61 IGSAQGDILGGSILGFDVAVGIGLILAGAYLDIKALDGITKKAFOFTWKGVMILAGVV 120
Db 143 V-----GGGLILSVDALENGILL-----QYNAEYQVSAGESLSPGCI 180

QY 121 T-----MAVTRLTEILPFTFANSYNRKLNVALGGFEPFDVAMQSSALG 170
Db 181 SGTYVGYIKDESIOKLDD-----NKMLDNOQNANLD-----KETLNQNKLD 223

QY 171 FELSFKKS 178
Db 224 YSVSFDKN 231
```

Search completed: March 14, 2003, 14:27:02
Job time : 28.47 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:19:56 ; Search time 30.8258 Seconds
(without alignments)
765.116 Million cell updates/sec

Title: US-09-508-487-23
 Perfect score: 908
 Sequence: 1 MNKELIFILVIFCATSSFAO.....DINMGOASALGELSEKKS 177

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002.*

1:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Query	Score	Query			DB	ID	Description
			Match	Length				
1	908	100.0	177	20	AAV04280	Borrelia garinii I		
2	786.5	86.6	178	20	AAV04279	Borrelia afzelii A		
3	737	81.2	179	20	AAV19808	B. burgdorferi ant		
4	737	81.2	179	20	AAV04278	Borrelia burgdorfe		
5	667.5	73.5	158	20	AAV19809	B. burgdorferi ant		
6	170.5	18.8	95	20	AAV04282	p11 protein sequen		
7	111	12.2	25	20	AAV04281	Borrelia burgdorfe		
8	88.5	9.7	451	22	AAU29340	Novel mar regulate		
9	88.5	9.7	451	22	AAG98374	Escherichia coli p		
10	87	9.6	413	22	AAB04097	Botulism toxin hea		

11	87	9.6	848	22	AAB504082	Botulinum toxin hea
12	87	9.6	852	23	AGG69077	Botulinum neurotox
13	87	9.6	858	19	AAW56018	Recombinant botuli
14	87	9.6	1070	21	AAV93308	A manganese supero
15	87	9.6	1095	21	AAV93311	A manganese supero
16	87	9.6	1169	19	AAW56017	Recombinant botuli
17	87	9.6	1291	19	AAW68392	Clostridium botuli
18	82	9.0	1456	22	AGC07111	Novel human diagno
19	82	9.0	1456	22	AGC24314	Novel human diagno
20	81.5	9.0	206	22	AAU35869	Helicobacter pylor
21	80	8.8	222	20	AAV34454	Porphorymonas ging
22	80	8.8	227	20	AAV34411	Porphorymonas ging
23	79.5	8.8	333	23	ABP27364	Streptococcus poly
24	79.5	8.8	726	22	AAU55177	Propionibacterium
25	79	8.7	479	20	AAV19665	B. burgdorferi ant
26	79	8.7	504	20	AAV19364	B. burgdorferi ant
27	78	8.6	398	22	AAV87898	C. glutamicum SRT
28	78	8.6	544	22	AAAG90173	C. glutamicum prote
29	78	8.6	544	22	AAV878997	C. glutamicum SRT
30	78	8.6	544	22	AAE19943	C. glutamicum SRT
31	77	8.5	243	23	ABP38116	Corynebacterium gl
32	76.5	8.4	345	22	AA876781	Staphylococcus epi
33	76.5	8.4	417	22	AG930916	Corynebacterium gl
34	76	8.4	237	22	AGB27217	C. glutamicum prote
35	76	8.4	283	18	AAW55463	Novel human diagno
36	76	8.4	293	18	AAW55244	H. pylori ORF 05ae
37	76	8.4	456	22	ABG58293	H. pylori ORF 05ae
38	76	8.4	781	22	ABG26754	Drosophila melanog
39	75.5	8.3	392	22	AAU67004	Novel human diagno
40	75.5	8.3	904	21	AAV51119	Propionibacterium
41	75.5	8.3	935	15	AAV55552	E. coli 0-157 vero
42	75.5	8.3	935	15	AAV55550	EHEC eae gene prod
43	75	8.3	215	19	AAW98808	Amino acid sequenc
44	75	8.3	474	23	AAE16782	H. pylori GHPO 132
45	75	8.3	998	18	AAW33621	Human transporter
46	75	8.3	998	18	AAW33621	Elmella tenella 45

ALIGNMENTS

RESULT 1	
AA04280	
ID	AA04280 standard; Protein; 177 AA.
XX	
XX	AA04280;
XX	
DT	17-JUN-1999 (first entry)
XX	
DE	Borrelia garinii IP90 protein.
XX	
XX	Borrelia; p13 antigen; spirochaete;
KW	
XX	
OS	Borrelia garinii.
XX	
XX	WC0912960-A2.
PN	
XX	
PD	18-MAR-1999.
XX	
XX	
PF	04-SEP-1998; 98WO-1B01424.
XX	
PR	16-SEP-1997; 97US-0059036.
XX	
PR	10-SEP-1997; 97DK-0001041.
XX	
XX	(SYMB-) SYMBICOM AB.
PA	
XX	
XX	Bergstroem S;
PI	
XX	
XX	WPI; 1999-215027/18.
DR	
DR	N-PSDB; AAX30100.
XX	
XX	Nucleic acid from Borrelia burgdorferi
PT	
PT	protein p13
XX	
XX	

PT caused by Borrelia, particularly Lyme disease

PS Claim 12; Page 72; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.

XX Sequence 179 AA;

Query Match 81.2%; Score 737; DB 20; Length 179;
Best Local Similarity 81.0%; Pred. No. 2.7e-73;

Matches 145; Conservative 14; Mismatches 18; Indels 2; Gaps 2;

QY 1 MNKELIIFILVIFCAFSSFAQ-DDSKS-TFNLGAGEKFLVYETNKKDSLVPPLNLFGLFGF 58

Db 1 MNKLLIFVLTFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPPFLNLFGLFGF 60

59 IGSAQGDILGGSILGLFDVAGIGLITGAYLDIKDFDNNAKKADFKWTGCKGMVLGAVV 118

Db 61 IGSAQGDILGGSILGLFDVAGIGLITGAYLDIKALDGIITKKAQFQWTGKGMVLGAVV 120

QY 119 TMVTRLTEIIVLPFTFANNYNRKLNLSNIALGCFEFSFDINMGQASALGFLSFKKSY 177

Db 121 TMVTRLTEIILPFTFANSYNRKLKNSLNVALGCFEFSFDVAMGQSSALGFLSFKKSY 179

RESULT 4

AAV04278

ID AAV04278 standard; Protein; 179 AA.

XX AAY04278;

DT 17-JUN-1999 (first entry)

DE Borrelia burgdorferi B31 protein.

KW Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis.

OS Borrelia burgdorferi.

XX W09912960-A2.

XX 18-MAR-1999.

XX 04-SEP-1998; 98WO-1B01424.

PR 16-SEP-1997; 97US-0059036.

PR 10-SEP-1997; 97DK-0001041.

XX (SYMB-) SYMBICOM AB.

PA Bergstroem S;

PI WPI; 1999-215027/18.

DR N-PSDB; AAX30098.

XX Nucleic acid from Borrelia burgdorferi encoding virulence associated
PT protein P13

XX Claim 31; Page 108-109; 118pp; English.

XX The present sequence represents a Borrelia burgdorferi B313 protein.
CC The present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC Borrelia burgdorferi B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most

CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by B. burgdorferi sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections, in standard immunoassays or amplification/hybridization
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within B. burgdorferi
CC sensu lato, but is absent from other Borrelia species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,
CC compared with use of whole bacteria as antigen.

XX Sequence 179 AA;

Query Match 81.2%; Score 737; DB 20; Length 179;
Best Local Similarity 81.0%; Pred. No. 2.7e-73;

Matches 145; Conservative 14; Mismatches 18; Indels 2; Gaps 2;

QY 1 MNKELIIFILVIFCAFSSFAQ-DDSKS-TFNLGAGEKFLVYETNKKDSLVPPLNLFGLFGF 58

Db 1 MNKLLIFVLTFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPPFLNLFGLFGF 60

QY 59 IGSAQGDILGGSILGLFDVAGIGLITGAYLDIKDFDNNAKKADFKWTGCKGMVLGAVV 118

Db 61 IGSAQGDILGGSILGLFDVAGIGLITGAYLDIKALDGIITKKAQFQWTGKGMVLGAVV 120

QY 119 TMVTRLTEIIVLPFTFANNYNRKLNLSNIALGCFEFSFDINMGQASALGFLSFKKSY 177

Db 121 TMVTRLTEIILPFTFANSYNRKLKNSLNVALGCFEFSFDVAMGQSSALGFLSFKKSY 179

RESULT 5

AAV19809

ID AAV19809 standard; Protein; 158 AA.

XX AAY19809;

DT 19-JUL-1999 (first entry)

DE B. burgdorferi antigenic protein, t933.aa.

KW Antigenic protein; vaccine; Lyme disease; infection; detection.

OS Borrelia burgdorferi.

XX W09859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

PR 03-SEP-1997; 97US-0057483.

PR 20-JUN-1997; 97US-0050359.

PR 22-JUL-1997; 97US-0053344.

PR 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (MEDI-) MEDIMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
PI WPI; 1999-189980/16.

DR N-PSDB; AAX61506.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases
PT caused by Borrelia, particularly Lyme disease

XX Claim 12; Page 73; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in

CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the *Borrelia* genus. The products can also
CC be used for detection of members of the *Borrelia* genus.

SQ Sequence 158 AA;
Query Match 73.5%; Score 667.5; DB 20; Length 158;
Best Local Similarity 82.3%; Pred. No. 1.1e-65;
Matches 130; Conservative 13; Mismatches 14; Indels 1; Gaps 1;
QY 21 DSKS-TFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAQGDILGSLILGFDAY 79
Db 1 NDSKNGAFGWSAGEKLLVYETKQDPVLPFLNLFGLFGIGSFAQGDILGSLILGFDAY 60
QY 80 GTGLIUTGAYLDIKDFDNNAKADKWTGKGMMLAGVVTMAVTRTEITVLPTFANNYN 139
Db 61 GTGLIAGAYLDIKALDGTITKAAFWGTGKGMMLAGVVTMAVTRTEITVLPTFANNYN 120
140 RLKLSNLALGGFEPSPFDINMGQASALGFLSFKKSY 177
Db 121 RLKLSNLVALGGFEPSPFDVAMQSSALGFLSFKKSY 158

RESULT 6
AAY04282
ID AAY04282 standard; Protein; 95 AA.
XX
AC AAY04282;
XX
DT 17-JUN-1999 (first entry)
XX
DE p11 protein sequence.
XX
KW *Borrelia*; P13 antigen; spirochaete; vaccine; infection; diagnosis.
XX
OS Unidentified.
XX
PN WO9912960-A2.
XX
PD 18-MAR-1999.
XX
PF 04-SEP-1998; 98WO-IB01424.
XX
PR 16-SEP-1997; 97US-0059036.
PR 10-SEP-1997; 97DK-0001041.
XX (SYMB-) SYMBICOM AB.

PI Bergstroem S;
XX
DR WPI; 1999-215027/18.
DR N-PSDB; AAX30115.
XX
PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated protein P13
XX
PS Example 8; Page 118; 118pp; English.
XX
CC The present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most
CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections, in standard immunoassays or amplification/hybridization
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within *B. burgdorferi*
CC sensu lato, but is absent from other *Borrelia* species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,

CC compared with use of whole bacteria as antigen. The present sequence
CC represents a p11 protein sequence from a best hit search.

SQ Sequence 95 AA;
Query Match 18.8%; Score 170.5; DB 20; Length 95;
Best Local Similarity 45.0%; Pred. No. 4.6e-11;
Matches 36; Conservative 11; Mismatches 32; Indels 1; Gaps 1;
QY 1 MKKFLIFILVIFCAFSSFAODDSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIG 60
Db 1 MKKIFTLILIFGLTIEIFATKDTQNIKGI-ESFNKYDKKKNPIGPFLLNLFDPFGIG 59
QY 61 SFAQGDILGSLILGFDAY 80
Db 60 SFAQGDYIGGSGVILGPNLLG 79

RESULT 7
AAY04281
ID AAY04281 standard; peptide; 25 AA.
XX
AC AAY04281;
XX
DT 17-JUN-1999 (first entry)
XX
DE *Borrelia burgdorferi* B313 N-terminal peptide.
XX
KW *Borrelia*; P13 antigen; spirochaete; vaccine; infection; diagnosis.
XX
OS *Borrelia burgdorferi*.
XX
PN WO9912960-A2.
XX
PD 18-MAR-1999.
XX
PF 04-SEP-1998; 98WO-IB01424.
XX
PR 16-SEP-1997; 97US-0059036.
PR 10-SEP-1997; 97DK-0001041.
XX (SYMB-) SYMBICOM AB.
XX
PI Bergstroem S;
XX
DR WPI; 1999-215027/18.
XX
PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated protein P13
XX
PS Example 6; Page 49; 118pp; English.

CC The present sequence represents a *Borrelia burgdorferi* B313 N-terminal
CC peptide. The present invention describes an isolated nucleic acid (I)
CC that: (i) encodes a polypeptide fragment (II) immunologically reactive
CC with rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most
CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections, in standard immunoassays or amplification/hybridization
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within *B. burgdorferi*
CC sensu lato, but is absent from other *Borrelia* species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,
CC compared with use of whole bacteria as antigen.

SQ Sequence 25 AA;
Query Match 12.2%; Score 111; DB 20; Length 25;

ID AAG98374 standard; Protein; 451 AA.

ID	AAB04097 standard; Protein; 413 AA.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

THIS PAGE BLANK (USPTO)

Qy 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCATAATACAGTTTAATA 180
Db 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCATAATACAGTTTAATA 180
Qy 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Db 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Qy 241 GTTTTGTCTCAAGATGATTTTAAATCTTAAATCTGGGAGCGGAGAGAAAATTTTGG 300
Db 241 GTTTTGTCTCAAGATGATTTTAAATCTTAAATCTGGGAGCGGAGAGAAAATTTTGG 300
Qy 301 TTTATGAACATAATGAAGAAGATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Db 301 TTTATGAACATAATGAAGAAGATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Qy 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGAGGTTCTCTTATTTCTGGATTG 420
Db 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGAGGTTCTCTTATTTCTGGATTG 420
Qy 421 ATGCGGTTGGTATAGGTTTAAATCTTACAGGAGCTTATTTGGACATCAAGGATTTTGATA 480
Db 421 ATGCGGTTGGTATAGGTTTAAATCTTACAGGAGCTTATTTGGACATCAAGGATTTTGATA 480
Qy 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGGAATGATGTTGGCAGGTG 540
Db 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGGAATGATGTTGGCAGGTG 540
Qy 541 TGGTTACTATGGCTGTGACAAGATTTGACAGAAATTTGTTCCCAATTTACATTTGCTAATA 600
Db 541 TGGTTACTATGGCTGTGACAAGATTTGACAGAAATTTGTTCCCAATTTACATTTGCTAATA 600
Qy 601 ATTATAACAGGAAGCTGAAAAATAGTCTTAATATAGCCTTGGAGGATTTTTCAGCCTAGTT 660
Db 601 ATTATAACAGGAAGCTGAAAAATAGTCTTAATATAGCCTTGGAGGATTTTTCAGCCTAGTT 660
Qy 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Db 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Qy 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749
Db 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749

RESULT 3

US-09-508-487-22
; Sequence 22, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Borrelia garinii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (192)..(725)
US-09-508-487-22

Query Match 100.0%; Score 749; DB 19; Length 749;
Best Local Similarity 100.0%; Pred. No. 4.4e-122;

Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTTGAAATGCTAAATTTTATGGTTA 60
Db 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTTGAAATGCTAAATTTTATGGTTA 60
Qy 61 AATCAAGAAGCTCTATTGTTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
Db 61 AATCAAGAAGCTCTATTGTTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
Qy 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCATAATACAGTTTAATA 180
Db 121 AATAGTTCACAAAGCCCTTTTAAATTCATTAATATGCTACCATAATACAGTTTAATA 180
Qy 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Db 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAAATCTTTTGGCTTTTCTA 240
Qy 241 GTTTTGTCTCAAGATGATTTTAAATCTTAAATCTGGGAGCGGAGAGAAAATTTTGG 300
Db 241 GTTTTGTCTCAAGATGATTTTAAATCTTAAATCTGGGAGCGGAGAGAAAATTTTGG 300
Qy 301 TTTATGAACATAATGAAGAAGATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Db 301 TTTATGAACATAATGAAGAAGATCTCTGTGACCAATTTTATTTGAACCTTTTTTATAGGGT 360
Qy 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGAGGTTCTCTTATTTCTGGATTG 420
Db 361 TCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGAGGTTCTCTTATTTCTGGATTG 420
Qy 421 ATGCGGTTGGTATAGGTTTAAATCTTACAGGAGCTTATTTGGACATCAAGGATTTTGATA 480
Db 421 ATGCGGTTGGTATAGGTTTAAATCTTACAGGAGCTTATTTGGACATCAAGGATTTTGATA 480
Qy 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGGAATGATGTTGGCAGGTG 540
Db 481 ATAATGCTAAAAAGCTGATTTTAAAGTGGACTTGGGTTAAGGGAATGATGTTGGCAGGTG 540
Qy 541 TGGTTACTATGGCTGTGACAAGATTTGACAGAAATTTGTTCCCAATTTACATTTGCTAATA 600
Db 541 TGGTTACTATGGCTGTGACAAGATTTGACAGAAATTTGTTCCCAATTTACATTTGCTAATA 600
Qy 601 ATTATAACAGGAAGCTGAAAAATAGTCTTAATATAGCCTTGGAGGATTTTTCAGCCTAGTT 660
Db 601 ATTATAACAGGAAGCTGAAAAATAGTCTTAATATAGCCTTGGAGGATTTTTCAGCCTAGTT 660
Qy 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Db 661 TTGATATTAAACATGGGCCAAGCTAGTCTCTTGGTTTGGACTGCTCTTTCAAAAAAGCT 720
Qy 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749
Db 721 ATTAATTTTATTTATCTAGAAAATGGGTG 749

RESULT 4

US-09-153-447-20
; Sequence 20, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 20
; LENGTH: 862
; TYPE: DNA

```

; ORGANISM: Borrelia afzelii
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: (219)..(755)
US-09-153-447-20

```

Query Match 81.0%; Score 607; DB 15; Length 862;
Best Local Similarity 90.1%; Pred. No. 4.3e-97;
Matches 673; Conservative 0; Mismatches 70; Indels

Matches 673; **Conservative** 0; **Mismatches** 70; **Indels** 4; **Gaps** 2;

QY	1	TAGAAATTTTTCACAAATAAGATATATGTTTAAAAAGAAATGGAATTCGTAAATTTTATCGTTGA	60
Db	27	TAGAATTTTTCGACAAATAAAGACATATATAAAAAGAAATGGAATTCGTAAATTTTATCGTCA	86
QY	61	AATCAAGAAGCTCTATTGGTAAGCGAAATTCGAGTAAACAATTTGAAAAAGTTAAATTTA	120
Db	87	AATCAAGAAGCTCTATTGGGAAGCGAAATTTCAAGTAAATCTTTGAAAAAGTTAAATTTA	146
QY	121	AATAGTTTCCAAAAGCCCTTTTTAAATTTTCATTAATATGCTACCATAATACCAG-TTTAAAT	179
Db	147	AATAGTTTAAAAACCTTTTTTAAATTTTCATTAATATGTTACTATTAATACCAGTTTAAAT	206
QY	180	AAAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAATCTTTTGGTGCCTTTTCT	239
Db	207	AAAGAGGTTTTTATGAATAAATTTTAAATTTGTTTGGTTTGTCTACCCCTTTTGTGTTTCT	266
QY	240	AGTTTCTCTCAA--GATGAATCTAAAAGCACATTTTAACTCGGAGCGGAGAGAAATTT	296
Db	267	AGCTTTGCTCAGCTGATGAATCTTAAAGGCGCTTTTAAATTTGGAGCGGAGAGAAACCT	326
QY	297	TTGGTTTATGAAACTAATAAGAAGATCTCTTGTACCACTTTTATTTGAACCTTTTTTTA	356
Db	327	TTAGCTTATGAACACTAGTAAGAAGATCCTATTGTGCCATTTTATTTGAACCTTTTTTTA	386
QY	357	GGGTTCGGGATAGGTTCTTTTTTGTCTCAAGGAGATATCCTTTGGAGGTTCTCTTATCTTGA	416
Db	387	GGGTTTGGAAATAGGTTCTTTTGTCTCAAGGAGATATCTTTGGGGGTTTTCTTATCTTGA	446
QY	417	TTTTCATCGGTTGGTATAGGGTTAATACTTTACAGGAGCTTATTTGGACATCAAGGATTTT	476
Db	447	TTTGTACGAGTTGGTATAGGGTTAATACTTTACAGGAGCTTATTTAGATATCAAGAGCTTT	506
QY	477	GATAAATAATGCTATAAAAAGCTGATTTTAACTGGACATGGGGTAAGGGAATGATTTGGCA	536
Db	507	GATAAGAATGCTCCAAAAGCGCTTTTAACTGGACATGGGGTAAGGGAATGATTTGGCA	566
QY	537	GGTGTGGTTACTATGGCTGTGACAGATGACAGAAATTTGTTCTCCATTTACATTTGCT	596
Db	567	GGTGCAGTTTACTATGGCTGTGACAAGATGACAGAAATTTATATCCGTTTACATTTGCT	626
QY	597	AATAAATTAACAGGAAGCTGAAAAATAGTCTTAATATAGCCCTGGAGGATTTGAGCCT	656
Db	627	AATAGTTTAAATAGGAACATGAAAAATAGCCCTTAATATAGCTTTTGGAGGGTTTGAACCT	686
QY	657	AGTTTTGATATTAACTATGGCCCAAGCTAGTGCCTCTTGGTTTGGACTGCTTTCAAAAAA	716
Db	687	AGTTTTGATATTAAATGCGCCCAAGCTAGGCGCTCTTGGGTTTGAACATCTTTCAAAAA	746
QY	717	AGCTATTAATTTTATTTATCTAGAAA	743
Db	747	AGTTATTAATTTTATTTATTTATTTAA	773

RESULT 5

```

: RECD 1
: US-09-156-447-20
: Sequence 20, Application US/09156447
: GENERAL INFORMATION:
: APPLICANT: BERGSTROM, Sven
: TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
: FILE REFERENCE: 454312-3130.1
: CURRENT APPLICATION NUMBER: US/09/156.447
: CURRENT FILING DATE: 1998-09-15
: PRIOR APPLICATION NUMBER: 60/059,036

```

```

; PRIORITY DATE: 1997-09-10
; PRIORITY APPLICATION NUMBER: 1041/97
; PRIORITY FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Borrelia afzelii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (219)..(755)
US-09-156-447-20

```

Query Match
81.0%; Score 607; DB 15; Length 862;

Query Match 81.0%; Score 607; DB 15; Length 862;
Best Local Similarity 90.1%; Pred. NO. 4.3e-97;
Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

QY	1	TAGAAATTTTCAACAATAAAGATATATGTTAAAGAAATTTGAAATGCTAAATTTTATGCTGTTA	60
Db	27	TAGAAATTTTTCGACAAAATAAAGACATATATTAAGAAATTTGAAATTTGCTAAATTTTATGCTG	86
QY	61	AATCAAGAAGCTCTATTGGTAAAGCAATTTTCAGAGTAAACAAATTTGAAAAAGTTTAAATTTTA	120
Db	87	AATCAAGAAGCTCTATTGGAGCGGAATTTCAAGTAATATCTTTGAAAAAGTTTAAATTTTA	146
QY	121	AATAGTTTCCAAAAGCCCTTTTAAAAATTTCAATTAATATGCTACCATTAATACCAG-TTTAAAT	179
Db	147	AATAGTTTAAAAACCTTTTAAAAATTTCAATTAATATGTTACTATTAATACCAGCTTTTAAT	206
QY	180	AAAGGGTTTTTATGAATAAAGTTTTTAAATTTTATTTTGGTAATCTTTTGTGCTTTTCT	239
Db	207	AAAGAGGTTTTTATGAATAAATTTTAAATTTGTTGTTTGTGCTACCCCTTTTGTGTTTCT	266
QY	240	AGTTTTCCTCAA--GATGATTTCTAAAGCACATTTTAACTGGAGCGGAGAGAAAATTT	296
Db	267	AGCTTTGCTCAAGCTGATGATTTCTAAAGCGCTTTTAAATTTGGAGCGGAGAGAAAAC	326
QY	297	TTGGTTTTATGAAACTAATAAGAAGATTTCTTTGTACCAATTTTATTTGAACCTTTTTTTA	356
Db	327	TTAGCTTATCAAACTAGTAAGAAGATCCTATTGTGCCATTTTATTGAACCTTTTTTTA	386
QY	357	GGGTCGGGATAGGTTCTTTTGTCTCAAGGAGATATCCTTTGGAGGTTCTCTTATCTTTGA	416
Db	387	GGGTTTGGAAATAGGTTCTTTTGTCTCAAGGAGATATTTCTTTGGGGTTTTCTTATCTTTGA	446
QY	417	TTTGATCGGTTGGTATAGGGTTTAACTTTACAGGAGCTTATTTGGACATCAAGGATTTT	476
Db	447	TTTGATGCAGTTGGTATAGGGTTTAACTTTACAGGAGCTTATTAGATATCAAGCTCTT	506
QY	477	GATAATAATGCTAAAAAAGCTGATTTTAAAGTGGACTTTGGGGTAAGGGAATGATTTGGCA	536
Db	507	GATAAAGAAATGCTCCAAAAGCCGCTTTTAAAGTGGACTTTGGGGTAAGGGAATGATTTGGCA	566
QY	537	GGTGTGGTTACTATGCTGTGACAAAGATTCACAGAAATTTGTTCTTCCATTTACATTTGCT	596
Db	567	GGTGCAGTTTACTATGCTGTGACAAAGATTCACAGAAATTTATTTCCGTTTACATTTGCT	626
QY	597	AATAAATTAACAGGAAGCTGAAAAATAGTCTTTAATATAGCCCTTTGGAGGATTTGAGCCT	656
Db	627	AATAGTTATTAATAGGAACCTGAAAAATAGCCCTTAAATATAGCTTTTGAGGGTTTGAGCCT	686
QY	657	AGTTTTGATATTAACTAGGCCCCAAGCTAGTGTCTTTGGTTTTGGACTGTCTTTCAAAAAA	716
Db	687	AGTTTTGATATTAAATAGGCCCCAAGCTAGGCTCTTTGGTTTGAACACTATCTTTCAAAAAA	746
QY	717	AGCTATTAAATTTATTATCTAGAAAA743	
Db	747	AGTTTAAATTTTATTATTTATTTATAA773	

RESULT 6
US-09-508-487-20


```
; Sequence 20, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Borrelia afzelii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (219)..(755)
; US-09-508-487-20

Query Match      81.0%; Score 607; DB 19; Length 862;
Best Local Similarity 90.1%; Pred. No. 4.3e-97;
Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

QY 1 TAGAATTTTCAACAATAAAGATATTTGTTAAAGAAATGGAATGCTAAATTTATGGTTA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27 TAGAATTTTCAACAATAAAGACATTTAAAGAAATGGAATGCTAAATTTATGGTCA 86
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 AATCAAGAAGCTCTATTGGTAAGGAATTTTCGAGTACAAATTTGAAAAAGTTAAATTA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 AATCAAGAAGCTCTATTGGGAAGCAATTTCAAGTAATACCTTTGAAAAAGTTAAATTA 146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 AATAGTTCCAAAAGCCCTTTTAAATTTCAATTAATATGCTACCAATAATACCAG-TTTAAT 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 AATAGTTTAAAACCCCTTTTAAATTTCAATTAATATGCTACCAATAATACCAGTTTAAAT 206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 180 AAGGGGTTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAATTCCTTTTGGCTTTTCT 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 AAGAGGTTTTTATGAATAAATTTTAAATTTGTTTGTAGCTTTTGTGTTTTTCT 266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 240 AGTTTCTCTCAA---GATGATCTTAAAGCACCTTTTAATCTGGGAGCGGAGAAAAATTT 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 AGCTTTCTCAAGCTGATGATCTTAAAGGCGCTTTTAAATTTGGGAGCGGAGAAAAACTT 326
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 297 TTGTTTTATGAACTAATAAGAAATCTCTTGTACCAATTTTATTTGAACCTTTTTTTA 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 327 TTAGCTTATGAACTAGTAAGAAATCCATTTGTGCCATTTTATTTGAACCTTTTTTTA 386
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 357 GGGTTCGGGATAGTTCTTTTGTCTCAAGGAGATATCCCTTTGGAGTTCTCTTATTTCTTGGGA 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 GGGTTTTGGAATAGTTCTTTTGTCTCAAGGAGATATCTTGGGGGTTTTCTTATTTCTTGGGA 446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 417 TTTGATCCGTTGCTATAGGTTAAATCTTACAGGAGCTTATTTGGACATCAAGGATTTT 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 447 TTTGATCCGTTGCTATAGGTTAAATCTTACAGGAGCTTATTTAGATATCAAGGCTTT 506
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 477 GATTAATATGCTAAAAAGCTGATTTTAAAGTGGACTTTGGGTAAGGGAATGATCTTGGCA 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 507 GATAAGATGCTCCAAAGCGCTTTTAAAGTGGACTTTGGGTAAGGGAATGATCTTGGCA 566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 537 GGTGCGTTACTATGCGTGTGACAAGATTTGACAGAAATTTGTTCTTCATTTACATTTGCT 596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 567 GGTGCGTTACTATGCGTGTGACAAGATTTGACAGAAATTTATTTATTCGTTTACATTTGCT 626
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 597 AATAAATTTAACAGGAGCTGAAAAATAGTCTTAAATATAGCTTTGGGAGGATTTGAGCCT 656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 627 AATAGTTTAAATAGGAACCTGAAAAATAGGCTTTAAATATAGCTTTGGGAGGTTTGGAGCCT 686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 657 AGTTTTTGATATTAACATGGGCCAAGCTAGTGTCTTGTGTTTGGACTGTCTTTCAAAAAA 716
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 687 AGTTTTCATATAATATGGCCCAAGCTAGCGCTCTTGGGTTTGAACATATCTTTCAAAAAA 746
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 717 AGCTATTAAATTTTATTATCTAGAAAA 743
    || ||||| ||||| || |||
Db 747 AGTTATTAAATTTTATTATTTATTATAA 773
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-153-447-18
; Sequence 18, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 60/059,036
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 1041/97
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (170)..(709)
; US-09-153-447-18

Query Match      74.5%; Score 558.2; DB 15; Length 759;
Best Local Similarity 87.0%; Pred. No. 1.6e-88;
Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;

QY 24 ATTCTGTTAAAAGCAATGAAATGCTAAATTTTATGGTTAAATCAAGAAGCTCTATTGGTAAG 83
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ATTCTGTTAAAAGCAATGAAATGCTAAATTTTATGGTTCAATCAAGNAGCTCTATTGGGAAG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 84 CGAATTTTCGAGTAACAATTTGAAAAAGTTTAAATTTTAAATAGTTTCCAAAAGCCTTTTTTA 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CGAATTTCAAGCAATAATTTTGAAGAAAGTTTAAATTTTAAATTAACCTTTTAAAGAACTTTTA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 144 AATTTTCATTAATATGCTACCAATAATACCAG-TTTAATAAAGGGGTTTTTATGAATAAGTT 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AATTTTCATTAATATGCTACCAATAATACCAGTTTAAATAAAGGGGTTTTTATGAATAAAGTT 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 203 TTTAATTTTATTTTGGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATGATTC-T- 260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TTTAATTTTGTGTTTGGCAACCTTTTGTGTTTCTAGCTTTGCTCAAGCTAATGATTC 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 261 ----AAAAGCACTTTTAAATCTGGGAGCGGAGAAAAATTTTGGTTTATGAACATAATAA 316
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TAAAAATGGTTCGTTTGGGATGAGTCTGGAGAAAAACTTTTGGTTTATGAACACTAGCAA 300
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 317 GAAAGATCTCTTCTACCAATTTTATGAACCTTTTATAGGTTTCGGATAGGTTCTTT 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GCAAGATCCTTATGTTACCACTTTTATGAACCTTTTATAGGTTTGAATAGGCTCCTT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 377 TGCTCAAGGAGATATCTTGGAGTTCTCTTATTTCTTGGATTTGATCGGTTGGTATAGG 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TGCTCAAGGAGATATCTTGGAGTTCTCTTATTTCTTGGATTTGATCGGTTGGTATAGG 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 437 GTTAATCTTACAGGAGCTTATTTGGACATCAAGGATTTTGAATAATGCTAAAAAGC 496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 GCTTATACTTCGGGGGCTTATTTGGATATCAAGCGCTTGTATGCTATTACTAAAAAAGC 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 497 TGATTTTAAGTGGACTTGGGCTAGGGGAATGATGTTGGCAGGTGTTGCTACTATGGCTGT 556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 TGCCTTCAATAGGACTTGGGGTAAGGGAGTTATGTTAGCAGGTGTTGTTACTATGGCTGT 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 557 GACAAGATTGACAGAAATTTGTTCTTCCATTTACATTTGCTTAATATTAACAGGAAGCT 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 541 GACAAAGATTAAACAGAAATATTCTCCATTACATTTGCTAATAGTATATAATAGGAGCT 600
QY 617 GAAAAATAGTCTTAATATAGCCCTTGGGAGGATTGTAGCCTAGTCTTGTGATATTAACATGGG 676
Db 601 AAAAAATAGCCCTTAATGTAGCTTGTAGGAGATTGGAACCTAGTCTTGTGATGTTCGAATGGG 660
QY 677 CCAAGCTAGTGCCTCTGGTCTTGGAGTCTGTCTTCAAAAAAGCTATTAATTTATTATC 736
Db 661 CCAATCCAGTGCCTCTGGGTTTGAAGTCTCTTCAAAAAAGCTATTAATTTATTATT 720
QY 737 TAGAAAAATGGGTG 749
Db 721 AAAAAATGGGTG 733

RESULT 8
US-09-156-447-18
; Sequence 18, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTRÖM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (170)..(709)
US-09-156-447-18

Query Match 74.5%; Score 558.2; DB 15; Length 759;
Best Local Similarity 87.0%; Pred. No. 1.6e-88;
Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;

QY 24 ATGTGTTAAAGAAATTCGAAATTTGCTAATTTATGTTTAAATCAAGAAGCTCTATTGCTAAG 83
Db 1 ATGTGTTAAAGAAATTCGAAATTTGCTAATTTATGTTCAATCAAGAAGCTCTATTGCGAAG 60
QY 84 CGAATTTCCAGTAACAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCTTTTAA 143
Db 61 CGAATTTCAAGCAATTAATTTGAAAAAGTTAAATTTAAATAACTTTTAAAAACCTTTTAA 120
QY 144 AATTTCATTAAATATGCTACCAATAATACCAG-TTTAATAAAGGGTTTTTATGAATAAGTT 202
Db 121 AATTTCATTAAATATGCTACCAATAATACCAGTTTAAATAAAGGGTTTTTATGAATAAAGT 180
QY 203 TTTAATTTTATTTTGGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATCATCT-- 260
Db 181 TTTAATTTTGTGTCACCTTTTGTCTTTTCTAGTTTGTCTCAAGATCATCT-- 240
QY 377 TGTCTCAAGGAGATATCTTTGGAGGTCTCTTATCTTGTGATTTGATGCGGTTGGTATAGG 436
Db 361 TGTCTCAAGGAGATATCTTTGGAGGTCTCTTATCTTGTGATTTGATGCGGTTGGTATAGG 420
QY 437 GTTAATACCTACAGGAGCTTATTGGACATCAAGGATTTTGAATAATATGCTAAAAAAGC 496
```

```
Db 421 GCTTATACTTGCAGGGCTTATTGATATCAAAAGCGCTGATGGTATTACTAAAAAGC 480
QY 497 TGAATTTTAACTGACACTTGGGTAAAGGAATGATGTTGGCAGGTGCTGGTACTATGCTGT 556
Db 481 TGTCTTCATAGGACTTGGGTAAAGGAGTTATGTTAGCAGGTGCTGTTACTATGCTGT 540
QY 557 GACAAGATTGACAGAAATTTGCTTCCATTTACATTTGCTTAATAATATTAACAGGAGCT 616
Db 541 GACAAGATTAAACAGAAATTTATCTTCCATTTACATTTGCTTAATAATAGGAGCT 600
QY 617 GAAAAATAGTCTTAATATAGCCCTTGGGAGGATTGTAGCCCTAGTCTTGTGATTAACATGGG 676
Db 601 AAAAAATAGCCCTTAATGTAGCTTGTAGGAGATTGGAACCTAGTCTTGTGATTAACATGGG 660
QY 677 CCAAGCTAGTGCCTCTGGTCTTGGAGTCTGTCTTCAAAAAAGCTATTAATTTATTATC 736
Db 661 CCAATCCAGTGCCTCTGGGTTTGAAGTCTCTTCAAAAAAGCTATTAATTTATTATT 720
QY 737 TAGAAAAATGGGTG 749
Db 721 AAAAAATGGGTG 733

RESULT 9
US-09-508-487-18
; Sequence 18, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTRÖM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Borrelia burgdorferi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (170)..(709)
US-09-508-487-18

Query Match 74.5%; Score 558.2; DB 19; Length 759;
Best Local Similarity 87.0%; Pred. No. 1.6e-88;
Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;

QY 24 ATGTGTTAAAGAAATTCGAAATTTGCTAATTTATGTTTAAATCAAGAAGCTCTATTGCTAAG 83
Db 1 ATGTGTTAAAGAAATTCGAAATTTGCTAATTTATGTTCAATCAAGAAGCTCTATTGCGAAG 60
QY 84 CGAATTTCCAGTAACAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCTTTTAA 143
Db 61 CGAATTTCAAGCAATTAATTTGAAAAAGTTAAATTTAAATAACTTTTAAAAACCTTTTAA 120
QY 144 AATTTCATTAAATATGCTACCAATAATACCAG-TTTAATAAAGGGTTTTTATGAATAAGTT 202
Db 121 AATTTCATTAAATATGCTACCAATAATACCAGTTTAAATAAAGGGTTTTTATGAATAAAGT 180
QY 203 TTTAATTTTATTTTGGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATCATCT-- 260
Db 181 TTTAATTTTGTGTCACCTTTTGTCTTTTCTAGTTTGTCTCAAGATCATCT-- 240
QY 261 ----AAAAGCAGCTTTTAAATCTGGGAGCGGAGAAAAATTTTGGTTTATGAAGCTAATAA 316
Db 241 TAAAAATGCTGGTGTGGGATGAGTCTGAGAAAAAAGCTTTTGGTTTATGAAGCTAGCAA 300
```

QY 317 GAAAGATTCCTGTGACCATTCTTTATGAACTCTTTTATAGGGTTCGGGATAGGTTCTTT 376
Db 301 CGAAGATCTATGTACCATTCTTTATGAACTCTTTTATAGGGTTCGGGATAGGTTCTTT 360
QY 377 TGCTCAAGGAGATATCCTTGGAGGTTCTCTTATTTCTTGGATTTGATCGGTTGGTATAGG 436
Db 361 TGCTCAAGGAGATATCCTTGGAGGTTCTCTTATTTCTTGGATTTGATCGGTTGGTATAGG 420
QY 437 GTTAACTACTACAGAGCTTATTTGGACATCAAGGATTTTGCATATATATGCTAAAAAGC 496
Db 421 GCTTATACCTTCGGGGGCTTATTTGGATATCAAGCGCTTGTATGCTATTAATAAAGC 480
QY 497 TGATTTAAGTGGACTTGGGCTAAGGGAATGATGTTGGCAGGTGGTGTACTATGCGTGT 556
Db 481 TGCTTTCAATGAGCTTGGGCTAAGGGAATGATGTTGGCAGGTGGTGTACTATGCGTGT 540
QY 557 GACAAGATGACAGAAATGTTCTTCCATTTACATTTGCTAATAATAACAGGAAGCT 616
Db 541 GACAAGATTAACAGAAATATTTCTTCCATTTACATTTGCTAATAATAAGGAAGCT 600
QY 617 GAAAATAGTCTTAATATAGCTTGGGAGGATTTGAGGCTAGTTTGGATATTAACATGG 676
Db 601 AAAAAATAGCTTAATGATGCTTTAGGAGGATTTGAACTAGTTTGGATATTAACATGG 660
QY 677 CCAAGCTAGTCTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 736
Db 661 CCAATCAGTCTCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
QY 737 TAGAAATGGGTG 749
Db 721 ACNAAATGGGTG 733

RESULT 10

PCT-US98-12764-63/c
; Sequence 63, Application PC/TUS9812764
; GENERAL INFORMATION:
; APPLICANT: Gil Choi et. al.
; TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 155
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM: Diskette, 3.50 inch, 1.4Mb storage
; MEDIUM TYPE: HP Vectra 486/33
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/12764
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB370PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1761 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

PCT-US98-12764-63

Query Match 13.1%; Score 98.4; DB 1; Length 1761;
Best Local Similarity 49.7%; Pred. No. 1.4e-07;
Matches 336; Conservative 0; Mismatches 331; Indels 9; Gaps 3;
QY 1 TAGAATTTTCAACAAATAAAGATATTTGTTAAAAAGAAATGGAATTTGCTAATTTTATGGTTA 60
Db 1031 TGGCATACAAAAAATAGATAGCATAAATTTCAAAATTTAAAGAGATTTAAAGAAAACITTT 972
QY 61 AATCAAGAGCTCTATTGGTAAGCGAATTTTCGAGTAACAATTTGAAAAAGTTAAATTTA 120
Db 971 ACATTTAAATCTGCTTAGAACAGCAAGAAATTTCAACAATATCTTTTAAAGAAAGATTCA 912
QY 121 AATAGTTCCAAAGCCTTTTAAATTTCAATTAATCTACCATTAATACCAAGTTTAATA 180
Db 911 ACTCCATAAATAAATCTAAAAATTTGCTTAACATTTGTTTATCAAACTTTAATTTATTAG 852
QY 181 AAGG---GGTTTTATGAATAGTTTTTAAATTTTATTTTGGTAATCTTTTGTGCTTTT 237
Db 851 GAGGTAATTAATATGAAAAAATTTTCACATTAATAATTTTGTAGTTTAAACAATGC 792
QY 238 CTAGTTTTGCTCAAGATGATCTTAAAGCACTTTTAACTCGGAGCGGAGAGAAAATTTT 297
Db 791 AAATCTTTGCACAGATAAACTTGAAAAAGGTGTTGGAGT-GGAGATATCGCGACTGTA 733
QY 298 TGGTTTATGAACATAATAGAAAGATCTCTGTGTACCATTTTATTTGAACCTTTTATTAG 357
Db 732 TGAATAATGAAGCCGAAAAAGCAACCACTACTAGCACCATTTACTTTTGAATATCTTTTGT 673
QY 358 GGTTCGGGATAGTTCTTTTGTCTCAAGGAGATATCTCTTGGAGGTCTCTTATTTCTTGGAT 417
Db 672 CTTTAGGGATAGGATCTTTTGTCTCAAGGAGATATATTGGTGTGTGCTACTTGGCT 613
QY 418 TTGATCGCGGTGGTATAGGGTTAACTTACAGGAGCTTTATTGGACATCAAGGATTTTG 477
Db 612 CTCAGTGTCTGGAAG-----AATCTTATTAATGCTGGACACATGATCGTAGGAGGTA 558
QY 478 ATAATAATGCTAAAAAGCTGATTTTAACTGGACCTTGGGTAAAGGAATGATGTTGGCAG 537
Db 557 TTGGAGGCTTACAGAAAAGCAGCAAGTAAATACCGGAGGAGTATTATCAGGAATAG 498
QY 538 GTGTGTTACTATGCTGTGCAAGATTTGACAGAAATTTGTTCTCCATTTACATTTGCTA 597
Db 497 GAGGCTCACAATTTAGCATCTTACATAACTGGAATTTATTATTTCAATTTAAATTTGCTA 438
QY 598 ATAATAATAACAGGAGCTGAAAAATAGTCTTTAATAATAGCCCTTGGAGGATTTGAGCCTA 657
Db 437 ATAGATACAACGACGACCTTAAAAAAGGCTCGGCATTGCACTTGGCGGGGTTGAACCCA 378
QY 658 GTTTTGATATTAAACAT 673
Db 377 ATTTGACATTTGGAAT 362

RESULT 11

US-09-830-228-63/c
; Sequence 63, Application US/09830228
; GENERAL INFORMATION:
; APPLICANT: Gil Choi et. al.
; TITLE OF INVENTION: Borrelia burgdorferi Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 155
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM: Diskette, 3.50 inch, 1.4Mb storage
; MEDIUM TYPE: HP Vectra 486/33
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2

Qy 491 AAAAGCTGATTTTAAGTGGACTTGGGTAAGGGAATGATGTTGGCAGGTGTGTTACTAT 550
Db 3459 AACTGGGCATAT-----AATAACAACGATAGGAGTAGCAGCAT 3497
Qy 551 GGCTGTGACACAGATTGACAGAAATGTTCTTCCATTTCATTTGCTAATAATTATACAG 610
Db 3498 TATAGCATCCACACATAGCCTCACTTATTTCCATTTCATTTGCAATAAACAATGC 3557
Qy 611 GAAGCTGAAAAATAGTCTTAATATAGCCTTGGGAGGATTTGAGCCTAGTTTTTGATATTAA 670
Db 3558 AATCTTAAAAAAGACTCGGCATGATATGCGGGTTTGAACCCCAATTTTGATATTGG 3617
Qy 671 CATGGC 677
Db 3618 AATAAGC 3624

Search completed: March 15, 2003, 09:01:10
Job time : 1473.7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 14, 2003, 22:48:52 ; Search time 56.3755 Seconds
(without alignments)
10731.744 Million cell updates/sec

Title: US-09-508-487-20

Perfect score: 862

Sequence: 1 gattttcattggtccag.....atatatttctctgtctaa 862

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	6.4	416	10	US-09-960-352-4584
2	54.2	6.3	2000	9	US-09-938-842A-4130
3	51.8	6.0	640681	10	US-09-790-988-1
4	49.2	5.7	2000	9	US-09-938-842A-5006
5	49.2	5.7	2002	10	US-09-887-576-4
6	48.8	5.7	2000	9	US-09-938-842A-3659
7	48.4	5.6	416	10	US-09-960-352-4584
8	48.2	5.6	127197	9	US-09-754-853A-1
9	48	5.6	411	10	US-09-960-352-14521
10	47.4	5.5	419	10	US-09-960-352-11234
11	47.4	5.5	640681	10	US-09-790-988-1
12	47.2	5.5	2251	10	US-09-796-256A-11
13	47	5.5	1058	10	US-09-804-682-30
14	46.6	5.4	431	10	US-09-960-352-5558
15	46.6	5.4	2000	9	US-09-938-842A-4502
16	46.2	5.4	480	10	US-09-960-352-5301
17	46	5.3	479	10	US-09-960-352-12872
18	45.8	5.3	425	10	US-09-960-352-1496
19	45.4	5.3	446	10	US-09-960-352-3400

c 20	45.2	5.2	2000	9	US-09-938-842A-4747
c 21	45.2	5.2	5006	10	US-09-837-751-7
c 22	45	5.2	376	10	US-09-960-352-5087
c 23	44.8	5.2	516	10	US-09-960-352-5785
c 24	44.6	5.2	428	10	US-09-969-373-1172
c 25	44.6	5.2	513509	9	US-09-754-853A-4
c 26	44.4	5.2	413	10	US-09-960-352-2919
c 27	44.2	5.1	2846	9	US-10-008-016-1
c 28	44.2	5.1	4157	7	US-08-556-422-1
c 29	44	5.1	408	10	US-09-960-352-1221
c 30	43.8	5.1	696	10	US-09-770-693-9
c 31	43.8	5.1	73308	10	US-09-954-456-2276
c 32	43.6	5.1	371	10	US-09-969-373-1447
c 33	43.6	5.1	414	10	US-09-960-352-6528
c 34	43.6	5.1	419	10	US-09-960-352-11234
c 35	43.4	5.0	2172	9	US-09-966-880A-15
c 36	43.4	5.0	2818	9	US-09-966-880A-7
c 37	43.4	5.0	6564	9	US-09-966-880A-10
c 38	43.4	5.0	11204	9	US-09-966-880A-35
c 39	43.2	5.0	299	10	US-09-960-352-13947
c 40	43.2	5.0	385	10	US-09-960-352-1739
c 41	43.2	5.0	26048	10	US-09-764-869-1556
c 42	43	5.0	1367	9	US-09-938-842A-4999
c 43	43	5.0	9412	9	US-10-032-393-51
c 44	43	5.0	9412	9	US-10-032-393-52
c 45	43	5.0	9417	9	US-10-032-393-48

ALIGNMENTS

RESULT 1

US-09-960-352-4584
; Sequence 4584, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4584
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11
US-09-960-352-4584

Query Match 6.4%; Score 55; DB 10; Length 416;
Best Local Similarity 49.5%; Pred. No. 0.31;
Matches 142; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY	21	AATTTGTGAGATTTTCGACAAATAAGACATTATTAAAGAATTGAAATTTGCTAATTTTA 80
DB	129	AATTTTATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 188
QY	81	TGGTCAATCAAGAAGCTCTATTGGGAGCGAATTTCAAGTAATCTTTGAAAAAAGTTA 140
DB	189	TTTAAAAATTAATAATATAAATAATTAATAAATAAATAAATAAATAAATAAATAAATA 248
QY	141	AATTTAAATAGTTTTTAAACACCTTTTTTAAATTTTCATTAAATGTTACTATAATACCA 200
DB	249	AATTTTAAATTTTAAAGAAAAATTTTAAAAATTTTAAAAATTTTAAATTTATAGAACA 308
QY	201	TTTAATAAGAGCTTTTATGATGAATAATTTTAAATTTGTTTGTGCTACCCCTTTGTGT 260
DB	309	TATAAAAAAGATTTTATTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 368

Db 291 TGTGAGTACATGACATTTGTAATTAAGTCAAAATACAAAATTTATTTATTTAAACAAAA 350
Qy 164 TTTTAAATTTTCATTAAATATGTTACTATAATACCAGTTTTTAATAAGAGGTTTTTATGAA 223
Db 351 ACGTTTGTAGTAAGATAATATTTATTTTACAAAAAATTTGTTTAAATATTTGTTAAACA 410
Qy 224 TAAATTTTAAATGTTGTTTGTCTAGCCCTTTGTTGTTTCTAGCTTTGCTCAAGCTGA 283
Db 411 ATATGTTTATGCTAATTTTAAAGGATTCATTTGAAGTTTGTAGGATTTAAATTTGATTGA 470
Qy 284 TGATTTCTAAAAGCGCTTTTAAATTTGGGAGCGGGAGAAAAAATTTTAGCTTTATGAACCTAG 343
Db 471 TAATGTGATGACATTTTAAATTAAGTTTAAAGTAAATTAATTTATTTTGGAAAAAA 530
Qy 344 TAAGAAAGAT 353
Db 531 AATTAATGCT 540

RESULT 5
US-09-887-576-4
Sequence 4, Application US/09887576
Patent No. US2002014047A1
GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360 001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2002
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-4

Query Match 5.7%; Score 49.2; DB 10; Length 2002;
Best Local Similarity 47.4%; Pred. No. 3.5;
Matches 147; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
Qy 44 AAAGACATTTATAAAGAAATGCAATTTGCTAATTTATGTCACAAATCAAGAAGCTCTATT 103
Db 230 AAAGCAAGGATTTAGAAGATTTAAATCTTTAATTAATTAATTTTCAATTAATAATTAAT 289
Qy 104 GGAAGCAATTTCAAGTAATACITTTGAAAAAGTTAAATTAATAGTTTAAATAAGCT 163
Db 290 TGTGAGTACATGACATTTGTAATTAAGTTCAAAATACAAAAATTTATTTATTTAAACAAA 349
Qy 164 TTTTAAATTTTCATTAAATATGTTACTATAATACCAGTTTTTAATAAGAGGTTTTTATGAA 223
Db 350 ACGTTTGTAGTAAGATAATATTTATTTACAAAAAATTTGTTTAAATATTTGTTAAACA 409
Qy 224 TAAATTTTAAATGTTGTTTGTCTAGCCCTTTGTTGTTTCTAGCTTTGCTCAAGCTGA 283
Db 410 ATATGTTTATGCTAATTTTAAAGGATTCATTTGAAGTTTGTAGGATTTAAATTTGATTGA 469
Qy 284 TGATTTCTAAAAGCGCTTTTAAATTTGGGAGCGGGAGAAAAAATTTTAGCTTTATGAACCTAG 343
Db 470 TAATGTGATGACATTTTAAATTAAGTTTAAAGTAAATTAATTTATTTTGGAAAAAA 529

Qy 344 TAAGAAAGAT 353
Db 530 AATTAATGCT 539
RESULT 6
US-09-938-842A-3659
; Sequence 3659, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINI
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3659
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3659
Query Match 5.7%; Score 48.8; DB 9; Length 2000;
Best Local Similarity 46.7%; Pred. No. 4.1;
Matches 155; Conservative 0; Mismatches 177; Indels 0; Gaps 0;
Qy 44 AAAGACATTTATAAAGAAATGCAATTTGCTAATTTTATGTCACAAATCAAGAAGCTCTATT 103
Db 1486 AAAGCAAGGATTTAGAAGATTTAAATCTTTAATTAATTAATTTTCAATTAATAATTA 1545
Qy 104 GGAAGCAATTTCAAGTAATACITTTGAAAAAGTTAAATTTAAATAGTTTAAATAAGCT 163
Db 1546 TGTGAGTACATGACATTTGTAATTAAGTTCAAAATACAAAAATTTATTTATTTAAACAAA 1605
Qy 164 TTTTAAATTTTCATTAAATATGTTACTATAATACCAGTTTTTAATAAGAGGTTTTTATGAA 223
Db 1606 ACGTTTGTAGTAAGATAATATTTATTTAAACAAAAAATTTGTTTAAATATTTGTTAGCAA 1665
Qy 224 TAAATTTTAAATGTTGTTTGTCTAGCCCTTTGTTGTTTCTAGCTTTGCTCAAGCTGA 283
Db 1666 ATATATTTATGCTAATTTTAAAGGATTTAATTTGAAGTTTGTGGGATTTAAATTTGATTGA 1725
Qy 284 TGATTTCTAAAAGCGCTTTTAAATTTGGGAGCGGGAGAAAAAATTTTAGCTTTATGAACCTAG 343
Db 1726 TAATGTGATGACATTTTAAATTCAAAAATTTAAAGTAAATTAATTTATTTTGGAAAAAAG 1785
Qy 344 TAAGAAAGATCCTATTGTCCTCAATTTTATTGA 375
Db 1786 ATTTAATGCTAATGCGCATTTAGATTGTAATAA 1817
RESULT 7
US-09-960-352-4584/C
; Sequence 4584, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Machialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION A
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C

RESULT 10

US-09-960-352-11234
; Sequence 11234, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byett, John C.
; APPLICANT: Machalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11234
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 48-LIB3058-052-Q1-K1-D8
US-09-960-352-11234

Query Match 5.5%; Score 47.4; DB 10; Length 419;
Best Local Similarity 48.7%; Pred. No. 6.1;
Matches 129; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

Qy 3 TTTTTCATGGATCCAGAAATTTGTAGAAATTTTCGACAAATAAGACAAATTTATAAGAA 62
Db 146 TTTTTCATGGATCCAGAAATTTGTAGAAATTTTCGACAAATAAGACAAATTTATAAGAA 205
Qy 63 TTGAATTTGCTAAATTTATGCTCAATCAAGAGCTCTATTGGGAGCGAATTTCAAGTA 122
Db 206 TTAATTTTAAATTTTAAATTAATAAATAAATTTTATATATATATTTTAA 265
Qy 123 ATACTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTAAATTTCAATATA 182
Db 266 TTTAAATTAATTTTAAATTTTAAATTTTAAATTAATAATTTAAATAATAATAATA 325
Qy 183 TGTACTATATACCAAGTTTAAATAAGAGGTTTATGAATAAATTTTAAATTTGTTT 242
Db 326 AAATTTTAAAAAATTTTAAAAAATATTTTAAAAAATATTTTAAAAAATATTTTATAATAATAAT 385
Qy 243 TTGCTAGCCTTTTGTGTTTTTCTA 267
Db 386 TTTATTTTAAATATATATTTA 410

RESULT 11

US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790.988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 5.5%; Score 47.4; DB 10; Length 640681;

Best Local Similarity 51.7%; Pred. No. 11;
Matches 108; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
Qy 27 TAGAATTTTTCGACAAATAAGACATATTTAAAGAAATTTGAATTCCTTAATTTTATGGTCA 86
Db 625163 TATAAATTTGTATAGATAGACCTTTTGTAGAAAAACGAGAATAACTTTGGATTGATAT 625104
Qy 87 AATCAAGAAGCTCTATTGGGAGCGAATTTCAAGTAATTAATTTTGAAGAAAGTTAAATTTA 146
Db 625103 AGTTCATTAATTTTGTGATAAACTTTGGACATAAACTTTGGATAAAAAATCAATAAA 625044
Qy 147 AATAGTTTAAAAACCTTTTAAATTTCAATTAATATGTACTATAATACCAGTTTAAAT 206
Db 625043 GAAAAATGCCAAGAAGATATACCAATTTTATTAATTTATATACATAATTCCTCTCCTGTA 624984
Qy 207 AAAGAGGTTTTTATGAATAAATTTTAAAT 235
Db 624983 CGAGAAGCTGTAATTTACTAGTATTTTCAT 624955

RESULT 12

US-09-796-256A-11
; Sequence 11, Application US/09796256A
; Patent No. US20020078477A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Vincent L
; APPLICANT: Caraway, Daniel T
; APPLICANT: Smeltzer, Richard H
; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
; FILE REFERENCE: 50617/C-3532.0
; CURRENT APPLICATION NUMBER: US/09/796.256A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US60/033381
; PRIOR FILING DATE: 1996-12-16
; PRIOR APPLICATION NUMBER: 08/991677
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2251
; TYPE: DNA
; ORGANISM: Pinus taeda
US-09-796-256A-11

Query Match 5.5%; Score 47.2; DB 10; Length 2251;
Best Local Similarity 49.6%; Pred. No. 7.7;
Matches 121; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

Qy 50 ATTATTAAGAAATTTGAATTTGCTAAATTTTATGGTCAATCAAGAGCTCTATTGGGAAG 109
Db 619 ACTATAAATTTGAGTTGTCGATGAGTGTATGAGATTTTAAAGATGTTAAATTTATATA 678
Qy 110 CGAATTTCAAGTAATACCTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTATA 169
Db 679 TGTAAATTAATTTTATTTTGAATTAACAAAAATTAATAATGGATAAAAAATTTGTTGT 738
Qy 170 AATTTCAATTAATGCTTACTATAATACCAGTTTAAATAAAGAGGTTTATGAATAAAT 229
Db 739 AAATTTAGAGTAAATAATTTCAAAATCTAAATAATTAACACTATTTATTTTAAATAAT 798
Qy 230 TTTAATTTGTTTGTGCTAGCCTTTTGTCTTTTCTAGCTTTGCTCAAGCTGATGATC 289
Db 799 TGTGGTAAATTTTATCTATATATTTAAGTTAAAAATTTAGAAAAAATTAATTTAAATA 858
Qy 290 TAAA 293
Db 859 TAAA 862

RESULT 13

US-09-804-682-30/c
; Sequence 30, Application US/09804682
; Patent No. US20020106765A1

GENERAL INFORMATION:
APPLICANT: Kinders, Robert
APPLICANT: Corey, Michael J.
TITLE OF INVENTION: PAL-18 POLYPEPTIDES, NUCLEIC ACIDS
TITLE OF INVENTION: ENCODING THE SAME AND METHODS FOR SCREENING FOR OR
TITLE OF INVENTION: MODULATING THE SAME
FILE REFERENCE: 130001.406
CURRENT APPLICATION NUMBER: US/09/804,682
CURRENT FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 174
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 1058
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7, 9, 88, 462, 505, 531, 582, 586, 607, 624, 658, 679, 680,
LOCATION: 687, 700, 702, 707, 719, 724, 730, 736, 737, 742, 747, 748,
LOCATION: 756, 761, 764, 765, 767, 768, 769, 775, 776, 779, 781, 782,
LOCATION: 785, 787, 788, 792, 793, 794, 796, 799, 800, 807, 817
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 818, 819, 821, 822, 823, 828, 829, 830, 832, 837, 838, 849,
LOCATION: 850, 851, 853, 855, 856, 860, 861, 862, 863, 864, 871, 872,
LOCATION: 873, 879, 881, 883, 886, 888, 891, 892, 893, 896, 900, 903,
LOCATION: 904, 905, 906, 909, 913, 914, 916, 919, 921, 925, 929
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 930, 931, 950, 951, 954, 957, 963, 965, 966, 967, 968, 969,
LOCATION: 973, 980, 981, 984, 988, 990, 991, 994, 996, 998, 999, 1001,
LOCATION: 1006, 1008, 1009, 1010, 1017, 1020, 1021, 1023, 1024, 1025,
LOCATION: 1027, 1031, 1034, 1035, 1039, 1042, 1043, 1044, 1045
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 1046, 1058
OTHER INFORMATION: n = A,T,C or G
US-09-804-682-30

Query Match 5.5%; Score 47; DB 10; Length 1058;
Best Local Similarity 42.6%; Pred. No. 7.8;
Matches 191; Conservative 0; Mismatches 257; Indels 0; Gaps 0;
QY 25 TGTAGATTTTCGACAAATAAGACACATATTAAGAAGATTAAGTAAATGCTAAATTTATGTT 84
DB 732 TGNAAATNTTAGGTTATTTGGAACAAANGAAGAAATATTTTCNAAGAAGNAAAGGG 673
QY 85 CAATCAAGAGCTCTATTTGGGAGCGAATTCAGTAATACATTTTGAANAAGTTAAAT 144
DB 672 GAATCCAATTAACATNTATTCGTATAGAGAAAGAAAGAAATGTAGANTAGGACAGTT 613
QY 145 TAAATAGTTTAAACACCTTTTAAATTCATTAATATGTTACTATATACCAGTTTAA 204
DB 612 TTAANCACTGATGGCCCGCTTTTANATANTCTTGATGTATATGTTGAATACATGTCA 553
QY 205 ATAAAGAGTTTTATGAATTAATTTTAAATGTTGTTTTCGTAGCCTTTTGTGTTTTT 264
DB 552 AGGATATGTAATGTAATTTTCANITTTTAAATATTTTATGCAGTATATTTTCAATTTT 493
QY 265 CTAGCTTTGCTCAAGCTGATGATCTTAAAGCGCTTTTAAATTTGGGAGCGGAGAAAC 324
DB 492 GTGTTTATGTCAGGTTTGAAGAACTTCAAGNAGATATTAATAATGACACAGATAAAT 433
QY 325 TTTTAGCTTATGAACCTAGTAAGAAAGATCCTATTGTGCGCATTTTATTTGAACCTTTT 384
DB 432 TCATTTTTCAGAACTTATTTGCACTGCTTAGTTTATTTTCTTCTTCCCATCT 373
QY 385 TAGGGTTGGAATGAGTTCTTTTTCGTCAGGAGATATCTTGGGGTTTTCCTTATCTTG 444
DB 372 TAAGTGCATGCAATTTGTTGATGTTTAAATTAATAAACTTGACATATATTTTATTATGT 313
QY 445 GATTGATGATGTTGTTAGGCTTAAAT 472
|||||

DB 312 AATTGAATTCATTTATTTGCCCTTTTAAAT 285
RESULT 14
US-09-960-352-5558
Sequence 5558, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 5558
LENGTH: 431
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 24-LIB3057-024-Q1-K1-F7
US-09-960-352-5558
Query Match 5.4%; Score 46.6; DB 10; Length 431;
Best Local Similarity 46.0%; Pred. No. 8.4;
Matches 157; Conservative 0; Mismatches 184; Indels 0; Gaps 0;
QY 124 TACTTGTGAAAAAGTTAAATTTAAATAGTTTTTAAAAACCTTTTAAATTTCAATTAAT 183
DB 57 TATTATAATTTGCTTATCTATATAAATTTATTTATTTTATTTTATTTTATTTTATTTT 116
QY 184 GTTACTATATAACCAGTTTTTAAATAAGAGGCTTTTATGAATAAATTTTAAATTTGTTCTT 243
DB 117 ATTTTATTTTATATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 176
QY 244 TGCTAGCCTTTTGTGTTTTTCTAGCTTTGCTCAAGCTGATGATCTTAAAGCGCTTTTA 303
DB 177 TTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 236
QY 304 ATTGGAGCGGGAGAAAACCTTTTAGCTTATGAACACTAGTAAGAAAGATCCTTATGTCG 363
DB 237 ATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 296
QY 364 CATTTTATTTGAACCTTTTATTTAGGTTTGGAAATAGTCTTTTCTCAAGGAGATATTC 423
DB 297 TTTATTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 356
QY 424 TTGGGGGTTTCTTATTTCTTTGGATTTTGATGACAGTTGGTATA 464
DB 357 TATTATTTTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 397
RESULT 15
US-09-938-842A-4502/c
Sequence 4502, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRP1300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4502
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4502

Query Match 5.4%; Score 46.6; DB 9; Length 2000;
Best Local Similarity 50.7%; Pred. No. 9.6; Indels 0; Gaps 0;
Matches 112; Conservative 0; Mismatches 109;
Qy 43 TAAAGACATTATTAAGAAGATTGAAATGCTAATTTATGCTCAATCAAGAAGCTCTAT 102
Db 1742 TAAAACTTTAGTAGATCTTTTGACTCTTGTATTCAATCAAAACTTTATTAATTTCTGT 1683
Qy 103 TGGGAAGCGAATTTCAAGTAATACTTTGAAAAAGTTAAATTAATAGTTTTAAAAACC 162
Db 1682 TAAAAATGATTATTAGTTTTTTGATTGTGAATTTAAAGTCTTTTAAAAATTTTCATAAATC 1623
Qy 163 TTTTTTAAATTTTCATTATATGTTACTATATACCAAGTTTAAATAAGAGGTTTTTATGA 222
b 1622 TGAATTTTATGAATCAAAATTTAATAAGTCAATAAATTTTGTGTGTTTAAATTAA 1563
Qy 223 ATAAATTTTAAATGTTGTTTGTAGCTTTTGTGTTTTT 263
Db 1562 AAAATGTATGATGTTAATGAATTAACATATGTTGTTATT 1522

Search completed: March 15, 2003, 14:21:27
Job time : 525.376 secs

US PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:19:56 ; Search time 31 Seconds
(without alignments)
765.116 Million cell updates/sec

Title: US-09-508-487-21

Perfect score: 904

Sequence: 1 MNRFLIVLLAFVCFSSFAQ.....DINMGASALGFELSPKSKY 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	904	100.0	178	AA04279	Borrelia afzelii A
2	786.5	87.0	177	AA04280	Borrelia garinii I
3	769.5	85.1	179	AA19808	B. burgdorferi ant
4	769.5	85.1	179	AA04278	Borrelia burgdorfe
5	687.5	76.1	158	AA19809	B. burgdorferi ant
6	160	17.7	95	AA04282	p11 protein sequen
7	124	13.7	25	AA04281	Borrelia burgdorfe
8	97	10.7	345	AA076781	Corynebacterium gl
9	97	10.7	417	AA093096	C glutamicum prote
10	95.5	10.6	479	AA19965	B. burgdorferi ant

11	95.5	10.6	504	20	AA19964	B. burgdorferi ant
12	85	9.4	364	23	AB50049	Listeria monocytog
13	83.5	9.2	572	21	AA01595	Streptococcus pneu
14	83.5	9.2	650	22	AA03748	Streptococcus pneu
15	82.5	9.1	215	19	AA09808	H. pylori GHPO 132
16	81.5	9.0	259	22	AA041896	Human polyepitide
17	81.5	9.0	413	22	AA040097	Botulism toxin hea
18	81.5	9.0	848	22	AA04082	Botulism toxin hea
19	81.5	9.0	852	23	ABG69077	Botulinum neurotox
20	81.5	9.0	858	19	AA056018	Recombinant botuli
21	81.5	9.0	1070	21	AA093308	A manganese supero
22	81.5	9.0	1095	21	AA093311	A manganese supero
23	81.5	9.0	1169	19	AA056017	Recombinant botuli
24	81.5	9.0	1291	19	AA068392	Clostridium botuli
25	81	9.0	305	19	AA071476	Helicobacter polyp
26	79	8.7	251	22	AA04855	Micromonospora eve
27	78.5	8.7	663	22	AA034477	E. coli cellular p
28	78.5	8.7	663	22	AA038487	Salmonella typhi c
29	78.5	8.7	663	22	AA098862	E. coli growth and
30	77.5	8.6	444	16	AA079372	Rhizobium DMP. Rh
31	77	8.5	490	23	AB054849	Lactococcus lactis
32	76.5	8.5	571	22	AA058659	Propionibacterium
33	76.5	8.5	635	23	AB054275	Lactococcus lactis
34	76	8.4	411	22	AA034379	Staphylococcus aur
35	76	8.4	432	22	AA037465	Staphylococcus aur
36	75.5	8.4	206	22	AA035859	Helicobacter pylor
37	75.5	8.4	382	23	AB047392	Listeria monocytog
38	75.5	8.4	463	23	AB049960	Listeria monocytog
39	75	8.3	534	18	AA032302	Arabidopsis thalia
40	75	8.3	534	18	AA036137	A. thaliana inorga
41	75	8.3	534	19	AA046748	Arabidopsis thalia
42	75	8.3	534	23	AB092038	Herbicidally activ
43	75	8.3	535	23	AB092665	Herbicidally activ
44	75	8.3	538	19	AA046749	Tomato phosphate t
45	74.5	8.2	180	22	AB066943	Drosophila melanog

ALIGNMENTS

RESULT 1
AA04279
ID AA04279 standard; Protein: 178 AA.
AC AA04279;
XX
XX 17-JUN-1999 (first entry)
DT
XX
XX Borrelia afzelii ACAI protein.
DE
XX
XX Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis.
KW
XX
XX Borrelia afzelii.
OS
XX
XX WO9912960-A2.
PN
XX
XX 18-MAR-1999.
PD
XX
XX 04-SEP-1998; 98WO-IB01424.
PF
XX
XX 16-SEP-1997; 97US-0059036.
PR
XX
XX 10-SEP-1997; 97DK-0001041.
XX
XX (SYMB-) SYMBICOM AB.
PA
XX
XX Bergstroem S;
PI
XX
XX WPI; 1999-215027/18.
DR
XX
XX N-PSDB; AAX30099.
DR
XX
XX Nucleic acid from Borrelia burgdorferi encoding virulence associated
PT
XX
XX protein P13

PS Claim 31; Page 111-112; 118pp; English.

XX The present sequence represents a Borrelia afzelii ACAI protein. The

CC present invention describes an isolated nucleic acid (I) that: (i)

CC encodes a polypeptide fragment (II) immunologically reactive with

CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of

CC Borrelia burgdorferi B313, but not reactive with most proteins from

CC other spirochaetes; and/or (ii) hybridizes under stringent conditions

CC to specific nucleic acid sequences but not to genomic DNA from most

CC other spirochaetes. (I), (II) and transformed cells are useful in

CC vaccines to protect against infection by B. burgdorferi sensu lato.

CC (I), (II) and antibodies raised against (II) are used to diagnose such

CC infections, in standard immunoassays or amplification/hybridization

CC tests. (I) are also used to produce recombinant (II). The 13 kD protein

CC is involved in virulence and is highly conserved within B. burgdorferi

CC sensu lato, but is absent from other Borrelia species (e.g. those

CC responsible for relapsing fever or avian borreliosis). It should provide

CC a more specific and sensitive antibody response, and diagnosis,

CC compared with use of whole bacteria as antigen.

XX

SQ Sequence 178 AA;

Query Match 100.0%; Score 904; DB 20; Length 178;

Best Local Similarity 100.0%; Pred. No. 2.3e-92;

Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKFLIWLIAFCVFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 60

DB 1 MNKFLIWLIAFCVFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 60

QY 61 GSFAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 120

DB 61 GSFAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 120

QY 121 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 178

DB 121 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 178

RESULT 2

AAV04280

ID AAY04280 standard; Protein; 177 AA.

XX

AC AAY04280;

XX

DT 17-JUN-1999 (first entry)

XX

Borrelia garinii IP90 protein.

XX

Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis.

XX

Borrelia garinii.

XX

OS

PN WO912960-A2.

XX

PD 18-MAR-1999.

XX

PF 04-SEP-1998; 98WO-IB01424.

XX

PR 16-SEP-1997; 97US-0059036.

XX

PR 10-SEP-1997; 97DK-0001041.

XX

(SYMB-) SYMBICOM AB.

XX

PA Bergstroem S;

XX

PI WPI; 1999-215027/18.

XX

DR N-PSDB; AAX30100.

XX

Nucleic acid from Borrelia burgdorferi encoding virulence associated

PT protein P13

XX

Claim 31; Page 114-115; 118pp; English.

XX The present sequence represents a Borrelia garinii IP90 protein. The

CC present invention describes an isolated nucleic acid (I) that: (i)

CC encodes a polypeptide fragment (II) immunologically reactive with

CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of

CC Borrelia burgdorferi B313, but not reactive with most proteins from

CC other spirochaetes; and/or (ii) hybridizes under stringent conditions

CC to specific nucleic acid sequences but not to genomic DNA from most

CC other spirochaetes. (I), (II) and transformed cells are useful in

CC vaccines to protect against infection by B. burgdorferi sensu lato.

CC (I), (II) and antibodies raised against (II) are used to diagnose such

CC infections, in standard immunoassays or amplification/hybridization

CC tests. (I) are also used to produce recombinant (II). The 13 kD protein

CC is involved in virulence and is highly conserved within B. burgdorferi

CC sensu lato, but is absent from other Borrelia species (e.g. those

CC responsible for relapsing fever or avian borreliosis). It should provide

CC a more specific and sensitive antibody response, and diagnosis,

CC compared with use of whole bacteria as antigen.

XX

SQ Sequence 177 AA;

Query Match 87.0%; Score 786.5; DB 20; Length 177;

Best Local Similarity 86.5%; Pred. No. 2.6e-79;

Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIWLIAFCVFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 60

DB 1 MNKFLIWLIAFCVFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFGLFGFI 59

QY 61 GSFAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 120

DB 60 GSFAQGDILGGFLLGFDVAGVIGLIITGAYLDIKALDKNAPKAAFKWTWKGMMLAGAVT 119

QY 121 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 178

DB 120 MAVTRTEIIIPFTFANSYNRKLKNSLNTAFGGFEPFSDINMGQASALGFELSFKKSY 177

RESULT 3

AAV19808

ID AAY19808 standard; Protein; 179 AA.

XX

AC AAY19808;

XX

DT 19-JUL-1999 (first entry)

XX

B. burgdorferi antigenic protein, f933.aa.

XX

Antigenic protein; vaccine; Lyme disease; infection; detection.

XX

Borrelia burgdorferi.

OS

PN WO9859071-A1.

XX

PD 30-DEC-1998.

XX

PF 18-JUN-1998; 98WO-US12718.

XX

PR 03-SEP-1997; 97US-0057483.

XX

PR 20-JUN-1997; 97US-0050359.

XX

PR 22-JUL-1997; 97US-0053344.

XX

PR 22-JUL-1997; 97US-0053377.

XX

(HUMA-) HUMAN GENOME SCI INC.

XX

PA (MEDI-) MEDIMUNE INC.

XX

PI Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX

WPI; 1999-189980/16.

XX

DR N-PSDB; AAX61505.

XX

New isolated Borrelia burgdorferi nucleic acids - used to develop

PT products for the diagnosis, prevention and treatment of diseases

XX

PT caused by Borrelia, particularly Lyme disease

PS Claim 12; Page 72; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.

XX SQ Sequence 179 AA;

Query Match 85.1%; Score 769.5; DB 20; Length 179;
Best Local Similarity 84.9%; Pred. No. 2e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIYVLLAFVCFSSFAQADDSKS-AFNLGAGEKLLAYETSKQDPVFPFLNLFGLFGF 59

Db 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPVFPFLNLFGLFGF 60

QY 60 IGSPAQGDILGGFLLGLFDVAGIGLITGAYLDIKALDNAPKAAFKWTGKGMMLAGAV 119

Db 61 IGSPAQGDILGGSLILGFDVAGIGLILAGAYLDIKALDGITKKAFFQWTGKGMMLAGV 120

QY 120 TMATVRLTEIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKS 178

Db 121 TMATVRLTEIIPFTFANSYNRKLNLSNIALGFGFEPFDVAMGQSSALGFELSFKKS 179

RESULT 4

AAV04278

ID AAY04278 standard; Protein; 179 AA.

XX AC AAY04278;

XX DT 17-JUN-1999 (first entry)

XX DE Borrelia burgdorferi B31 protein.

XX KW Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis.

XX OS Borrelia burgdorferi.

XX PN W09912960-A2.

XX PD 18-MAR-1999.

XX PF 04-SEP-1998; 98WO-IB01424.

XX PR 16-SEP-1997; 97US-0059036.

XX PR 10-SEP-1997; 97DK-0001041.

XX PA (SYMB-) SYMBICOM AB.

XX PI Bergstroem S;

XX DR WPI; 1999-215027/18.

XX DR N-PSDB; AAX30098.

XX Nucleic acid from Borrelia burgdorferi encoding virulence associated

PT protein P13

XX Claim 31; Page 108-109; 118pp; English.

XX The present sequence represents a Borrelia burgdorferi B313 protein.
CC The present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC Borrelia burgdorferi B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most

CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by B. burgdorferi sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections, in standard immunoassays or amplification/hybridization
CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within B. burgdorferi
CC sensu lato, but is absent from other Borrelia species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,
CC compared with use of whole bacteria as antigen.

XX SQ Sequence 179 AA;

Query Match 85.1%; Score 769.5; DB 20; Length 179;
Best Local Similarity 84.9%; Pred. No. 2e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIYVLLAFVCFSSFAQADDSKS-AFNLGAGEKLLAYETSKQDPVFPFLNLFGLFGF 59

Db 1 MNKLLIFVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPVFPFLNLFGLFGF 60

QY 60 IGSPAQGDILGGFLLGLFDVAGIGLITGAYLDIKALDNAPKAAFKWTGKGMMLAGAV 119

Db 61 IGSPAQGDILGGSLILGFDVAGIGLILAGAYLDIKALDGITKKAFFQWTGKGMMLAGV 120

QY 120 TMATVRLTEIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKS 178

Db 121 TMATVRLTEIIPFTFANSYNRKLNLSNIALGFGFEPFDVAMGQSSALGFELSFKKS 179

RESULT 5

AAV19809

ID AAV19809 standard; Protein; 158 AA.

XX AC AAV19809;

XX DT 19-JUL-1999 (first entry)

XX DE B. burgdorferi antigenic protein, t933.aa.

XX KW Antigenic protein; vaccine; Lyme disease; infection; detection.

XX OS Borrelia burgdorferi.

XX PN W09859071-A1.

XX PD 30-DEC-1998.

XX PF 18-JUN-1998; 98WO-US12718.

XX PR 03-SEP-1997; 97US-0057483.

XX PR 20-JUN-1997; 97US-0050359.

XX PR 22-JUL-1997; 97US-0053344.

XX PR 22-JUL-1997; 97US-0053377.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (MEDI-) MEDIMUNE INC.

XX PI Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX DR WPI; 1999-189980/16.

XX DR N-PSDB; AAX61506.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases
PT caused by Borrelia, particularly Lyme disease

XX Claim 12; Page 73; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in

CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the *Borrelia* genus. The products can also
 CC be used for detection of members of the *Borrelia* genus.

XX Sequence 158 AA;

Query Match 76.1%; Score 687.5; DB 20; Length 158;
 Best Local Similarity 85.4%; Pred. No. 2.3e-68;
 Matches 135; Conservative 10; Mismatches 12; Indels 1; Gaps 1;

QY 22 DSKS-AFNLGAGEKLLAYETSKKDPVLPFLNLFLGFGTSGFAQGDILGGFLILGFDV 80

DB 1 NDSKAFGMSAGEKLLVETSKQDPVLPFLNLFLGFGTSGFAQGDILGGFLILGFDV 60

QY 81 GIGLITGAYLDIKALDKNAPKAAFKWTGKGMMLAGAVTMVTRTEIIPFTFANSYN 140

DB 61 GIGLILAGAYLDIKALDGTITKKAQFQWTGKGVMLAGVVTMAYTRTEIIPFTFANSYN 120

QY 141 RLKLSNLTAFFGFSPFDINMGQASALGFELSFKKS 178

DB 121 RLKLSNLVALGGFSPFDVAMGQSSALGFELSFKKS 158

RESULT 6

AA04282
 ID AAY04282 standard; Protein; 95 AA.

AC AAY04282;

DT 17-JUN-1999 (first entry)

XX pl1 protein sequence.

XX *Borrelia*; P13 antigen; spirochaete; vaccine; infection; diagnosis.

XX Unidentified.

XX WO9912960-A2.

XX 18-MAR-1999.

XX 04-SEP-1998; 98WO-IB01424.

XX 16-SEP-1997; 97US-0059036.

XX 10-SEP-1997; 97DK-0001041.

XX (SYMB-) SYMBICOM AB.

XX Bergstroem S;

XX WPI; 1999-215027/18.

XX N-PSDB; AAX30115.

PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
 PT protein P13

XX Example 8; Page 118; 118pp; English.

CC The present invention describes an isolated nucleic acid (I) that: (i)
 CC encodes a polypeptide fragment (II) immunologically reactive with
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within *B. burgdorferi*
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide
 CC a more specific and sensitive antibody response, and diagnosis,

CC compared with use of whole bacteria as antigen. The present sequence
 CC represents a pl1 protein sequence from a best hit search.

SQ Sequence 95 AA;

Query Match 17.7%; Score 160; DB 20; Length 95;

Best Local Similarity 44.4%; Pred. No. 5e-10;

Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;

QY 1 MNKFLIVLLAFVCSFAQADDSKSAFNLGAGEKLLAYETSKKDPVLPFLNLFLGFGI 60

DB 1 MKKIFLLIFLGLTIEIFA-TRDQNRKIEKGI-ESFNKYDKRKNPIGPFLNLFPGI 58

QY 61 GSFAQGDILGGFLILGFDV 81

DB 59 GSPVQGDYIGGSGVLGNLLG 79

RESULT 7

AA04281

ID AAY04281 standard; peptide; 25 AA.

XX AAY04281;

DT 17-JUN-1999 (first entry)

XX *Borrelia burgdorferi* B313 N-terminal peptide.

XX *Borrelia*; P13 antigen; spirochaete; vaccine; infection; diagnosis.

XX *Borrelia burgdorferi*.

XX WO9912960-A2.

XX 18-MAR-1999.

XX 04-SEP-1998; 98WO-IB01424.

XX 16-SEP-1997; 97US-0059036.

XX 10-SEP-1997; 97DK-0001041.

XX (SYMB-) SYMBICOM AB.

XX Bergstroem S;

XX WPI; 1999-215027/18.

PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
 PT protein P13

XX Example 6; Page 49; 118pp; English.

CC The present sequence represents a *Borrelia burgdorferi* B313 N-terminal
 CC peptide. The present invention describes an isolated nucleic acid (I)
 CC that: (i) encodes a polypeptide fragment (II) immunologically reactive
 CC with rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within *B. burgdorferi*
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide
 CC a more specific and sensitive antibody response, and diagnosis,
 CC compared with use of whole bacteria as antigen.

SQ Sequence 25 AA;

Query Match 13.7%; Score 124; DB 20; Length 25;

Best Local Similarity 96.0%; Pred. No. 9.5e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 41 TSKKDPVFPFLNLFGLFGIGSFAQ 65
Db 1 TSKQDPVFPFLNLFGLFGIGSFAQ 25

RESULT 8
AAB76781
ID AAB76781 standard; Protein; 345 AA.
XX AAB76781;
XX
DT 11-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MCT protein SEQ ID NO:544.
XX
KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
KW membrane construction and membrane transport protein; petroleum spill;
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;
KW identification; microorganism; fine chemical production; transformation;
KW genome mapping; genetic engineering.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100805-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB00926.
XX
PR 25-JUN-1999; 99US-0141031.
PR 08-JUL-1999; 99DE-1031454.
PR 08-JUL-1999; 99DE-1031478.
PR 08-JUL-1999; 99DE-1031563.
PR 09-JUL-1999; 99DE-1032122.
PR 09-JUL-1999; 99DE-1032124.
PR 09-JUL-1999; 99DE-1032125.
PR 09-JUL-1999; 99DE-1032128.
PR 09-JUL-1999; 99DE-1032180.
PR 09-JUL-1999; 99DE-1032182.
PR 09-JUL-1999; 99DE-1032190.
PR 09-JUL-1999; 99DE-1032191.
PR 09-JUL-1999; 99DE-1032209.
PR 09-JUL-1999; 99DE-1032212.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032228.
PR 09-JUL-1999; 99DE-1032229.
PR 09-JUL-1999; 99DE-1032230.
PR 14-JUL-1999; 99DE-1032927.
PR 14-JUL-1999; 99DE-1033005.
PR 14-JUL-1999; 99DE-1033006.
PR 27-AUG-1999; 99DE-1040764.
PR 27-AUG-1999; 99DE-1040765.
PR 27-AUG-1999; 99DE-1040766.
PR 27-AUG-1999; 99DE-1040830.
PR 27-AUG-1999; 99DE-1040831.
PR 27-AUG-1999; 99DE-1040832.
PR 27-AUG-1999; 99DE-1040833.
PR 31-AUG-1999; 99DE-1041378.
PR 31-AUG-1999; 99DE-1041379.
PR 31-AUG-1999; 99DE-1041395.
PR 03-SEP-1999; 99DE-1042077.
PR 03-SEP-1999; 99DE-1042078.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042088.
XX
PA (BADI) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX WPI; 2001-071486/08.
XX

DR N-PSDB; AAF68014.
XX
PT Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation -
XX
PS Claim 20; Page 925-926; 1119pp; English.
XX
CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAB76510 to
CC AAB76847. The MCT nucleic acids and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.
XX
SQ Sequence 345 AA;
XX
Query Match 10.7%; Score 97; DB 22; Length 345;
Best Local Similarity 23.1%; Pred. No. 0.025;
Matches 50; Conservative 24; Mismatches 62; Indels 80; Gaps 9;

QY 5 LIVVLLAFVCVFSFAQADDSKSAFNLGAGEKLLAYETSKK-----DPI 47
Db 63 MTALLLAFC-----LGLGMAVIKSDFLFKVTRERLVRVMTTITAFVPL 106
QY 48 VP-FLNLLFLGIGSFAQGDIL-----GGFLILGFDVIGLILTGAVLDIKALDKNA 100
Db 107 LPLFIFGIFLCMG-----NGGLEIMSAFCVKVLLA--VVGTLFLAIQFIAGAVSKN 160
QY 101 PKAAFKWTNGKMLAGAVTMVTRLTEIIPFTFANSYNRKLNLSNIAFGGFEP---- 156
Db 161 PWKLFE-----NMLPAYFTALGTSSAATIPVT----YQOTLKNDVDVNVAGFVPLCA 210
QY 157 -----SFDINMQASALGFEL 172
Db 211 TIHLASMMKIGLFTFAVFMVMEVGLSIGFL 246

RESULT 9
AAG93096
ID AAG93096 standard; Protein; 417 AA.
XX AAG93096;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 6850.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX

DR WPI: 2001-376931/40.
 XX N-PSDB; AAH68315.
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 XX
 PS Claim 17: SEQ ID NO: 6850; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Corynebacterium bacterium, and identifying a homologue of a gene derived
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX SQ Sequence 417 AA;
 Query Match 10.7%; Score 97; DB 22; Length 417;
 Best Local Similarity 23.1%; Pred. No. 0.032; Length 417;
 Matches 50; Conservative 24; Mismatches 62; Indels 80; Gaps 9;
 QY 5 LIVVLLAFVSSFAQADDSKSAFNLGAGKLLAYETSKK-----DPI 47
 Db 135 MTALLLAF-----LGLGMAVIKSDTLFKYTRLEVRVVMKTIATFVPL 178
 QY 48 VP-FLNLFLGFGIGSFAQGDIL-----GGPLILGFDVAGIGLITGAYLDIKALDKNA 100
 Db 179 LPLFTFGLFGMGM-----NGGLLEIMSAFGKVLILA--VVGPLFLAIQFIAGAVSKKN 232
 QY 101 PKAAFKWTGKGMMLAGAVTMAVTRLTEIIPFTFANSYNRKLNLSLATAFGGEP--- 156
 Db 233 PKWLFK-----NMLPAVETALCTSSAATIPVT---YQQLKRDVDVNVAGFVVPICA 282
 QY 157 -----SFDINNGQASALGFEL 172
 Db 283 TIHLAGSMNKIGLFTFAVVMYDMVGVLISIGFEL 318

RESULT 10
 AAY19965
 ID AAY19965 standard; Protein: 479 AA.
 AC AAY19965;
 XX

DT 19-JUL-1999 (first entry)
 DE B. burgdorferi antigenic protein, t843.aa.
 KW Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.
 OS

XX WO9859071-A1.
 XX

XX 30-DEC-1998.
 XX

XX 18-JUN-1998; 98WO-US12718.
 XX

XX 03-SEP-1997; 97US-0057483.
 XX

XX 20-JUN-1997; 97US-0050359.
 XX

XX 22-JUL-1997; 97US-0053344.
 XX

XX 22-JUL-1997; 97US-0053377.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.

PA (MEDI-) MEDIMMUNE INC.
 XX

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;
 XX

XX WPI: 1999-189980/16.
 DR

XX N-PSDB; AAX61662.
 XX

PT New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases
 PT caused by Borrelia, particularly Lyme disease
 XX

PS Claim 12: Page 139-140; 275pp; English.
 XX

CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus.
 XX

XX SQ Sequence 479 AA;
 Query Match 10.6%; Score 95.5; DB 20; Length 479;
 Best Local Similarity 25.5%; Pred. No. 0.056; Length 479;
 Matches 55; Conservative 25; Mismatches 67; Indels 69; Gaps 12;

QY 6 IVVLLAFVCF-----SSFAQADDSKSAFNLGA-----GEKLLAYETSKKDIPI--- 47
 Db 93 IVIILPFCVFLAIFFLIWSF---DDLIALSLGLFVFRPEYFGSFLYGLNLLPLGLCH 149

QY 48 ----VPF-----LLNLF-----LGFGLGSAQG--DILGFG-LILGFD 78
 Db 150 SILSPFEFTSLGGVEIVNGDVTVRGLKNIFYAQLDPSLGKFSFGAKISSGFYLSIMFG 209

QY 79 AVGIGLILTGAYLDIKALDKNAKPAFAKWTGKGMMLAGAVTMAVTRLTEIIPFTFANS 138
 Db 210 LPGAAL---GVYKGVHEDKNKVA-----LLFSGALTAFLTGITE---PLEFLFI 254

QY 139 YNRKLNLSLATAFGF-----EPSFDINNGQASALGF 170
 Db 255 FTAPLLYFVHAYSGFALLANFNFTVNTIGNSFSTGF 290

RESULT 11
 AAY19964
 ID AAY19964 standard; Protein: 504 AA.
 AC AAY19964;
 XX

DT 19-JUL-1999 (first entry)
 DE B. burgdorferi antigenic protein, f843.aa.
 KW Antigenic protein; vaccine; Lyme disease; infection; detection.

XX Borrelia burgdorferi.
 OS

XX WO9859071-A1.
 XX

XX 30-DEC-1998.
 XX

XX 18-JUN-1998; 98WO-US12718.
 XX

XX 03-SEP-1997; 97US-0057483.
 XX

XX 20-JUN-1997; 97US-0050359.
 XX

XX 22-JUL-1997; 97US-0053344.
 XX

XX 22-JUL-1997; 97US-0053377.
 XX

XX (HUMA-) HUMAN GENOME SCI INC.
 XX

XX (MEDI-) MEDIMMUNE INC.
 XX

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI: 1999-189980/16.
 DR N-PSDB; AAX81661.
 XX

PT New isolated *Borrelia burgdorferi* nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases
 PT caused by *Borrelia*, particularly Lyme disease
 XX

PS Claim 12: Page 139; 275pp; English.
 PS

CC This sequence represents a *Borrelia burgdorferi* (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the *Borrelia* genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the *Borrelia* genus. The products can also
 CC be used for detection of members of the *Borrelia* genus.
 XX

SQ Sequence 504 AA;
 Query Match 10.6%; Score 95.5; DB 20; Length 504;
 Best Local Similarity 25.5%; Pred. No. 0.059;
 Matches 55; Conservative 25; Mismatches 67; Indels 69; Gaps 12;

QY 6 IVLLAFVCF-----SFAQADSKSAFNILGA-----GEKLLAYETSKKDPI--- 47
 Db 118 IVILPCVFLAIFFLIWSF---DILIASLGLFVFRFEYFGSLYGLNRLLLPLGLH 174
 QY 48 ----VFP-----LLNLF-----LGFIGSFAQG--DILGGF-LILGFD 78
 Db 175 SILSPFPEFTSLGVEIVNGDVTVRGLKNIFYAQLDPSLCKFSGFAKISSGFYLSIMFG 234
 QY 79 AVGILGLTGTALDIKALDKNAPKAFKWTGKGMMLAGAVTMAVTRLEIIIPFTFANS 138
 Db 235 LPGAAL---GVYKIVHEDKKNVAA-----LLFSGALTAFLTGITE---PLEFLFI 279

QY 139 YNRKLNLSLTAFGF----EPSPDINMGQASALGF 170
 Db 280 FTAPLLVFAHYSGFALLANFENVVIGNSFSTGF 315

RESULT 12
 ABB50049
 ID ABB50049 standard; Protein; 364 AA.
 AC ABB50049;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE *Listeria* monocytogenes protein #2753.
 XX
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX

Listeria monocytogenes.
 OS
 XX
 PN WO200177335-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 11-APR-2001; 2001WO-FR01118.
 XX
 PR 11-APR-2000; 2000FR-0004629.
 XX
 PA (JNSP) INST PASTEUR.
 XX

Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia del Portillo F, Gomez-Lopez N;
 PI Maduenio E, De Pablos B, Wehlant J, Kaerst U, Entian K, Hauf J;

PI Rose M, Voss H;
 XX WPI: 2002-010914/01.
 DR

XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 PT and prevention of *Listeria* and related bacterial infections, and
 PT related polypeptides -
 PT
 XX

PS Claim 6; SEQ ID No 2754; 192pp; French.
 PS

CC The present invention relates to the genome sequence of *Listeria*
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytogenes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC

SQ Sequence 364 AA;
 Query Match 9.4%; Score 85; DB 23; Length 364;
 Best Local Similarity 23.5%; Pred. No. 0.59;
 Matches 32; Conservative 26; Mismatches 46; Indels 32; Gaps 6;

QY 37 LAYETSKDPIVPFLNLFLGFGISFAQGDILGGFLILGFDVAVGIGLILGTAYLDIKAL 96
 Db 71 IAYSIADRPGLAPGLMLGLIANSGAGPLGMLGGL-----VGYFVLIVKYLK--- 121
 QY 97 DKNAPKAAFKWTGKGM-----LAGAVTMAVTRLEIIIPFTFA-----NSYNRKLANS 146
 Db 122 ----PK-----WAQGLPMPIILISLVVGLLMFYVGVPIVWATEMTSFLQGMGS 171

QY 147 LNIAFG---GPEPSFD 159
 Db 172 MRFVGVGLGMAAFD 187

RESULT 13
 AAY81595
 ID AAY81595 standard; Protein; 572 AA.
 XX
 AC AAY81595;
 XX
 DT 24-MAY-2000 (first entry)
 XX
 DE *Streptococcus pneumoniae* type 4 protein sequence #95.
 XX
 KW *Streptococcus pneumoniae*; vaccine; screening; protein antigen;
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
 KW pneumococcal disease.
 XX

Streptococcus pneumoniae.
 OS
 XX
 PN WO200006737-A2.
 XX
 PD 10-FEB-2000.
 XX
 PF 27-JUL-1999; 99WO-GB02451.
 XX
 PR 27-JUL-1998; 98GB-0016337.
 PR 19-MAR-1999; 99US-0125164.
 XX


```
XX 29-JUL-1997; 97US-0902615.
PR 01-APR-1997; 97US-0833457.
PR 24-JUN-1997; 97US-0881227.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
PI WPI: 1998-542293/46.
DR N-PSDB; AAXI4527.
DR
XX
PT New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
XX
XX Claim 8; Page 1770-1771; 2054pp; English.
XX
XX This sequence represents a Helicobacter pylori GPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis.
XX
XX SQ Sequence 215 AA;
Query Match 9.1%; Score 82.5; DB 19; Length 215;
Best Local Similarity 23.0%; Pred. No. 0.57;
Matches 37; Conservative 18; Mismatches 41; Indels 65; Gaps 6;
QY 5 LIVVLLAFCVFSSFAQADDSKAFN-----LCAGEKLLAY-----ETS 42
DB 38 IVMYLMAFCIKVSYEDIKNKSAFNIFLLGSLTMAAGLKNVGFNFIGNAAQNFLA 97
QY 43 KKDPFVP--FLLNLF-----LGFIGSFAQGDILGGFLILGFDVAG 81
DB 98 HLDPLIAVLFIVALFYLSHYFFASITAHVSALFALFVGIGSHIQGVNLQELSLFLM 157
QY 82 IGLILT-----GAYLDIKALDKNAPKAAFKWTWG 110
DB 158 IMGILTPYGTGPSTIYYGSGYIQSKDF-----WKWG 188
```

Search completed: March 14, 2003, 14:24:24
Job time : 33 secs

THIS PAGE BLANK (USPTO)

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-161 <KLE>
A:Cross-references: GB:AE000790; NID:g2690224; PIDN:AAC66226.1; PID:g2690225; TIGR:BBA01
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 32.5%; Score 293.5; DB 2; Length 161;
Best Local Similarity 41.1%; Pred. No. 9.6e-20;
Matches 74; Conservative 24; Mismatches 61; Indels 21; Gaps 6;
QY 1 MNKFLIVLLAFCVSFSSFAQADSKSAFNLGAGEKLLAYETSKDPIVFPFLNLFGLFGI 60
DB 1 MKKIPTLILFGLTIQIFA-TRDQNRKEKGI-ESFNKYDKEKNPIGPFLLNLFPLFGI 58
QY 61 GSFAQGDILGGLILGDFAVGLITGAYLDIKALDKNAPKAFKWTGKGMMLAG-AV 119
DB 59 GSFVQGDYIGGSVGLGNLGAALMGTTILN-----HRETQLTGYILGVGA 106
QY 120 TMAVTRLT-EIIPPTFANSYNRKLNLSNIAFGGPEPFDINMGQASALGFELSFKSY 178
DB 107 SMVLISYTVSLIPTFANRHNENLKKLSAELAGFEPNEDLGIN-----GFQLSFKSY 161

RESULT 3
F70243
conserved hypothetical protein BBI31 - Lyme disease spirochete plasmid I/lp28-4
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: F70243
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70243
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-170 <KLE>
A:Cross-references: GB:AE000789; NID:g2690079; PIDN:AAC66194.1; PID:g2690091; TIGR:BBI31
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 31.5%; Score 285; DB 2; Length 170;
Best Local Similarity 38.8%; Pred. No. 6.1e-19;
Matches 69; Conservative 25; Mismatches 70; Indels 14; Gaps 3;
QY 1 MNKFLIVLLAFCVSFSSFAQADSKSAFNLGAGEKLLAYETSKDPIVFPFLNLFGLFGI 60
DB 7 MKKILTLILFSLTIQIFATQD--KLEKSVGSIETIMKYSEKATILAPFLNLFGLTGI 64
QY 61 GSFAQGDILGGLILGDFAVGLITGAYLDIKALDKNAPKAFKWTGKGMMLAGAVT 120
DB 65 GSFVQGDYIGGAVLGSLLGILCTAGNLTGHTDDETRA-----TTGHIITIGVCT 117
QY 121 MAVTRLTEIIPPTFANSYNRKLNLSNIAFGGPEPFDINMGQASALGFELSFKSY 178
DB 118 IIAASHIASLIIPPTFANKNANLKRGLDIAGFEPNEDIGIS-----GFQLSFKRY 170

RESULT 4
G70239
conserved hypothetical protein BBH41 - Lyme disease spirochete plasmid H/lp28-3
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: G70239
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: G70239
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-190 <KLE>
A:Cross-references: GB:AE000784; NID:g2690041; PIDN:AAC65989.1; PID:g2690045; TIGR:B
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 26.5%; Score 240; DB 2; Length 190;
Best Local Similarity 39.6%; Pred. No. 9.1e-15;
Matches 57; Conservative 23; Mismatches 52; Indels 12; Gaps 4;
QY 36 LLAYETSKK-DPIVFPFLNLFGLFGIGGFAQGDILGGLILGDFAVGIGLITGAYLDIK 94
DB 58 LLQYEKEQLIDLWIPVNLNLFPLFGVGVQGDYIGGCTLGFNLGLITLTGV-IQWK 116
QY 95 ALDKNAPKAFKWTGKGMMLAGAVTMAVTRLTEIIPPTFANSYNRKLNLSNIAFGGF 154
DB 117 NLKKEPASIS-----SMILLGMLTFGSSYLISIVLPVLFEDRYKKMLNRIIDELAGF 171
QY 155 EPSFDINMGQASALGFELSFKSY 178
DB 172 EPNLDIGMN-----GFQLSFKSY 190

RESULT 5
D70114
PI3 system, maltose and glucose-specific IABC component (malX) homolog - Lyme disease
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: D70114
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: D70114
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-574 <KLE>
A:Cross-references: GB:AE001124; GB:AE000783; NID:g2688003; PIDN:AAC66516.1; PID:g268
A:Experimental source: strain B31
C:Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosph
F.15-527/Domain: phosphotransferase system glucose-specific enzyme II, factor II homo

Query Match 10.6%; Score 95.5; DB 2; Length 574;
Best Local Similarity 25.5%; Pred. No. 0.49;
Matches 55; Conservative 25; Mismatches 67; Indels 69; Gaps 12;
QY 6 IIVLLAFCVF-----SSFAQADDSKSAFNLGA-----GKLLAYETSKKDP1--- 47
DB 188 IIVILPFCVFLAIFFLCIWSSF---DDLIASLGLFVRFYFGSGFLYGLNRLLLPLGLH 244
QY 48 ----VFP-----LLNLF-----LGFIGISFAQG--DILGFG-LILGFD 78
DB 245 SILSPFFETSLGGVEIVNGDVRGLKNIFYAQLDPSLGKSSGFAKISSGFYLSIMFG 304
QY 79 AVGIGLITGAYLDIKALDKNAPKAFKWTGKGMMLAGAVTMAVTRLTEIIPPTFANS 138
DB 305 LPGAAL---GVYKIGIVHEDKNKVA-----LLFSGALTAFTLTGITE---PLEFLFI 349
QY 139 YNRKLNLSNIAFGGF-----EPSFDINMGQASALGF 170

Qy 38 AYETSKKDPIVPFLNLFLFGIGSFAQDILGGFLILG---FDAVGIGLILTAYLDI 93

A;Map position: linear chromosome

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria species*.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1341
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00213.1; PID:g16411605; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2135

Query Match 9.4%; Score 85; DB 2; Length 364;
Best Local Similarity 23.5%; Pred. No. 2.8;
Matches 32; Conservative 26; Mismatches 46; Indels 32; Gaps 6;

Qy 37 LAYETSKDPVPELNLFLGFGSGSFAQGDILGFLILGFDVAVGIGLILTCAYLIDIKAL 96
Db 71 IAYSTADRPAGIAPGLMLGLIANSICAGFLGMLGGYL-----VGYFVLILVKYLVK--- 121
Qy 97 DKNAPKAAFKWTGKGM-----LAGAVTMAVTRLTEIIPFFA-----NSYNKKLKN 146
Db 122 ----PK-----WAQGLPMWMIIPLISSLVGLLMFYVGVPIVWATEAMTSFLQGMGS 171
Qy 147 LNIAGF---GPEPSFD 159
Db 172 MRFVFGAVLGAMAADF 187

RESULT 14

A41901
C:Species: Streptomyces glaucescens
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 22-Oct-1999
C:Accession: A41901; S27687
R:Guillou, P.G.; Hutchinson, C.R.
J. Bacteriol. 174, 3651-3658, 1992
A:Title: Sequence and transcriptional analysis of the Streptomyces glaucescens tcmAR tetA gene.
A:Reference number: A41901; MUID:92276347; PMID:1592819
A:Contents: GLA.O, ETH 22794
A:Accession: A41901
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-538 <GUI>
A:Cross-references: EMBL:M80674; NID:g153488; PIDN:AAA67509.1; PID:g153490
A:Note: sequence extracted from NCBI backbone (NCBIP:104630)
C:Superfamily: lincomycin-resistance protein lmrB
C:Keywords: transmembrane protein

Query Match 9.4%; Score 85; DB 2; Length 538;
Best Local Similarity 24.6%; Pred. No. 4.2;
Matches 33; Conservative 19; Mismatches 38; Indels 44; Gaps 6;

Qy 49 PFLNLFLGFGIG-----SFAQGDILGFLI--LGFD-----VGIGLITGAYLIDKA 95
Db 145 PCKLMAIGISGVVVGASTAAGPIIGLLVQHVGVAVFFINVPVGLAALVAGLILTD 204
Qy 96 LDKNAPKA-----AFKWTGK-----GVMLAGAV-----TMAVT 124
Db 205 RAERAPKSFVSGIVLLSGAMFCLVWGLIKAPMGWGLDRLTGLFLAAAVLAFAGFTLRES 264
Qy 125 RLTEIIPFFANS 138
Db 265 RATEPLMPLAMERS 278

RESULT 15

T19495

hypothetical protein C27A7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19495
R:Harris, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: T19132
A:Accession: T19495
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-674 <WIL>
A:Cross-references: EMBL:Z81041; PIDN:CA802785.1; GSPDB:GN00023; CESP:C27A7.3
A:Experimental source: clone C27A7
C:Genetics:
A:Gene: CESP:C27A7.3
A:Map position: 5
A:Introns: 61/1; 98/3; 220/3; 268/3; 340/1; 392/2; 475/2; 496/3; 537/1; 588/1

Query Match 9.3%; Score 84.5; DB 2; Length 674;
Best Local Similarity 24.0%; Pred. No. 5.8;
Matches 37; Conservative 21; Mismatches 47; Indels 49; Gaps 8;

Qy 4 FLIVVLLAFVCVFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVIVPFLNLFLGFGIGSF 63
Db 41 FLFILLICFSVSFSTLI-----IKNYQT-----FLISFFFRIVWSL 78
Qy 64 AOGDILGGLILGFDVAVGIGLILTCAYLIDIKALDKNAPKAAFKWTGKGMMLAGAVTMAV 123
Db 79 AMLSFIAGLVLLGLVAAA---TISGS-----KNLPTAEYKW-----AGCENLGK 119
Qy 124 TRLTE-----IIIPFT-FANSY-NRKLKNSLNI 149
Db 120 QCIDGFSTPPLVILSFDGFAKEYLERRIVKSL 153

Search completed: March 14, 2003, 14:25:41
Job time : 14.3333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:25:51 ; Search time 10.0562 seconds
(without alignments)
820.439 Million cell updates/sec

Title: US-09-508-487-19
Perfect score: 902
Sequence: 1 MNKLIFVLATCFVSSFAQ.....DVAMGSSALGFELSFKKSY 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published_Applications_AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	9.4	378	9 US-10-260-877-48	Sequence 48, Appl
2	83.5	9.3	384	9 US-09-738-626-3606	Sequence 3606, Ap
3	80.5	8.9	650	10 US-09-815-242-13341	Sequence 13341, A
4	80	8.9	487	9 US-09-738-626-4875	Sequence 4875, Ap
5	79.5	8.8	319	10 US-09-815-242-13924	Sequence 13924, A
6	79.5	8.8	417	9 US-09-738-626-6850	Sequence 6850, Ap
7	79	8.8	346	9 US-09-738-626-5811	Sequence 5811, Ap
8	78.5	8.7	513	9 US-09-738-626-6860	Sequence 6860, Ap
9	77.5	8.6	513	9 US-10-051-902-20	Sequence 20, Appl
10	77.5	8.6	513	9 US-10-051-909-20	Sequence 20, Appl
11	74.5	8.3	570	10 US-09-815-242-11951	Sequence 11951, A
12	74.5	8.3	708	9 US-09-738-626-5024	Sequence 5024, Ap
13	74	8.2	417	10 US-09-815-242-10165	Sequence 10165, A
14	73	8.1	451	10 US-09-741-669-422	Sequence 422, App
15	72.5	8.0	345	9 US-09-738-626-4062	Sequence 4062, Ap
16	72.5	8.0	468	10 US-09-815-242-5144	Sequence 5144, Ap
17	72.5	8.0	693	9 US-09-365-536A-14	Sequence 14, Appl
18	72.5	8.0	1161	8 US-08-910-386A-20	Sequence 20, Appl
19	72.5	8.0	1294	9 US-09-836-499-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-260-877-48

; Sequence 48, Application US/10260877

; Publication No. US20030021813A1

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Chovan, Linda E.

; APPLICANT: Hessler, Paul E.

; APPLICANT: Reich, Karl A.

; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME

; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF

; TITLE OF INVENTION: 'ESSENTIAL GENES'

; FILE REFERENCE: 6565 US.P1

; CURRENT APPLICATION NUMBER: US/10/260,877

; CURRENT FILING DATE: 2002-09-30

; PRIOR APPLICATION NUMBER: US/09/649,145

; PRIOR FILING DATE: 2000-08-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 48

; LENGTH: 378

; TYPE: PRT

; ORGANISM: H. influenzae

US-10-260-877-48

Query Match 9.4% Score 84.5; DB 9; Length 378;

Best Local Similarity 28.3%; Pred. No. 0.44; Mismatches 19; Gaps 5;

Matches 34; Conservative

QY 14 VFSSFAQANDSKNGAFMGSGEKLLVYETSQDPIVPP-----LLNLFLGFGIGSFAQG 67

Db 73 VKSFGLSGDGKTAIEMAAASGLHLVPEKRNPLTTSYGTGELIKLALDLGVESFILG 132

QY 68 DILGSSLILGPDVAVGILILAGAYLDIKALGQITKKAQFQWTKGVMLAGVVTMVTRL 127

Db 133 --IGGS---ATNDGGVGMLQA---LGMQCLDSQDKPIGF-----GGAELANIVKIDVOOL 179

RESULT 2

US-09-738-626-3606

; Sequence 3606, Application US/09738626

; Publication No. US20020197605A1

[illegible]

```
Db 253 IMIVAFAALFAWF-----IYQOSRAEK-----SGNDPLVPLEIFKRNFSLGNIC 297
QY 66 QGDILGSLILGFDVAGVIGLILAGAYLIDIKALDGTTRKKAQFQWTGKGVMLAGVVTMAVT 125
Db 298 -----IMAGFTVAGTPIIMLY-----FQAHGNNMAGFWMVPOA 335
QY 126 RLTEIILPFTFANSYNRKLKNSLNLVALGGFPEFDFVAMQSSALGF 171
Db 336 LMAAVLSPFV-----GKLVDRSN-----PGLMAALGF 362

RESULT 5
US-09-815-242-13924
; Sequence 13924, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13924
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13924

Query Match 8.8%; Score 79.5; DB 10; Length 319;
Best Local Similarity 27.5%; Pred. No. 1.2;
Matches 39; Conservative 23; Mismatches 39; Indels 41; Gaps 9;

QY 39 VYETSKDPI-----VPFLNLFLG--FGIGSFAQGDILGSLILGFDVAGVIGLILAGA 90
Db 138 LFATYFQEPVGNILSNIPFVGALESFAGFGLIAGVIL-----ALMIPTYRAV 188
QY 91 YLDI-KALDGTTRKKAQFQWTGKGVMLAGVVTMAVTRLTEILPFTFANSYNRKLKNSL 149
Db 189 MRDVFEGTPVMKESAY-----GI---GCTTWEV--IWRIVLPFT-----KNGV- 227
QY 150 VALGGFPEFDFVAMQSSALGF 171
Db 228 --IGGIMLGLRALGETMAVTF 247

RESULT 6
US-09-738-626-6850
; Sequence 6850, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
```

```
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6850
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6850

Query Match 8.8%; Score 79.5; DB 9; Length 417;
Best Local Similarity 24.7%; Pred. No. 1.6;
Matches 39; Conservative 22; Mismatches 50; Indels 47; Gaps 8;

QY 47 PIVP-FLNLFLGFGIGSFAQGDILGSLILGFDVAGVIGLILAGA----YLDIKAL--DG 99
Db 177 PLLPLFIPIGIFLGMG-----MNGGLEIMSAFGVLLAVVGTLLFLAIQIILAGA 227
QY 100 ITKKAQFQWTGKGVMLAGVVTMAVTRLTEILPFTFANSYNRKLKNSLNLVALGGFEP-- 157
Db 228 VSKKN--PWKLFKN-MLPAYFTALGTSSSAATIPVT---YQOTLKNDVDVNVAGFVVP 280
QY 158 -----SFDVAMQSSALGFEL 173
Db 281 CATIHLAGSMKIGLTFVAVVFMVMEVGVGLSIGFL 318

RESULT 7
US-09-738-626-5811
; Sequence 5811, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5811
```

```

Query Match      8.7%; Score 78.5; DB 9; Length 513;
Best Local Similarity 20.0%; Pred. No. 2.6;
Matches 38; Conservative 27; Mismatches 42; Indels 83; Gaps 9;

QY 49 VP-FLNLFLGFGIGSFAQ--GDILGGS�-ILGFDVAGIGLILAGAYLD----- 93
Db 17 VPAFLGIGITAVAGLUGMGRSVQGVIGGAIRKATLGLLIGAGATLVTASLEPLGAMINGAT 76
QY 94 -----IKALDQITKK-----AAQW----- 108
Db 77 GMRGVVPTNEAIGAGAEYGAQVAWLMLGFAISLVLAFTNLRVYVLTGHHVLFMSTM 136
QY 109 -----TW---GKGNVLACGVITMAVTRLTEIILPFTFANSYNRKLSNLVALGG 154
Db 137 LTILATAGFDAMIVVGVGALLLGMVS-----LP-AFAHPWTRRITGGDSIAIGH 187
QY 155 FEPSFDVAMG 164
Db 188 FGTAGYVAAG 197

RESULT 9

```

```
Query Match      8.6%; Score 77.5; DB 9; Length 513;  
Best Local Similarity 26.0%; pred.No. 3.3;  
Matches        Conservative 18; Mismatches   30; Indels    43; Gaps     5;  
  
QY       3 KLIFVLATCVTSPPAQANDSRNCAFCMSAGEKLLVYTSKODPIVPFLINLPLGFGIG 62  
          :|||:||::|||||::|::|::|:  
Db        61 EILGIULSLXLCFSFAGATSOR-----IGRLTVV----- 92  
          |||::|::|  
QY       63 SFAQGDLTGGSLIIGFDPAVGIGLIAGAYLDIKALDGITKAAPQTWKGVMLAGVVTM 122  
          |||::|::|
```

Db 93 -FAAVIFFVGSLLMGF-AVNYGMLMAGRFV---AGVG------GYGMIAPVYTA 137

Qy 123 AVT 125

Db 138 EIS 140

RESULT 11

US-09-815-242-11951
; Sequence 11951, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11951

; LENGTH: 570

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-815-242-11951

Query Match

Best Local Similarity 8.3%; Score 74.5; DB 10; Length 570;

Matches 39; Conservative 25; Mismatches 49; Indels 57; Gaps 8;

Qy 27 GAGMSAGEKLLVYETSKQDPVFPFLNLFGLFGGSGFAQ-----GDIL 70

Db 175 GAFAGVNLRLIVTG-----LHHLNUNVWFVFGFTDDETRGVVTGDLARFAGDPK 228

Qy 71 GGSILG-FDAVGICL-----ILAGAYLDIKA---LDGITKKAFFOWT 109

Db 229 GGQFMAGNFPWMFGLPAACLAAYNAPRRKRLIGLLSKALTAFLTGTVEPVEAFM 288

Qy 110 WKGVG-----MLAGVVTMAVTRTEITLPTTFANSYNRKLNLSNALG 153

Db 289 FLAPLLVLLHALLTG-LSMALTDLLDIRLGFTEG-----GAIDLALG 330

RESULT 12

US-09-738-626-5024

; Sequence 5024, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; PRIOR FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 5024

; LENGTH: 708

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-5024

Query Match 8.3%; Score 74.5; DB 9; Length 708;

Best Local Similarity 24.7%; Pred. No. 10;

Matches 38; Conservative 21; Mismatches 54; Indels 41; Gaps 8;

Qy 4 LLIFVLATFCVFSFAQANDSKNGAFG-----MSAGEKLLVYETSKQDPVFPFL----- 52

Db 532 ILVFIAATIQVRRSF-----YGHKWLQILVLGLTLTSLAISIVTITTFIGRLIA 581

Qy 53 --LNLFLGFGGSGFAQG--DILGSSLIG--FDA-----VGIGLILAG-----AYLDIK 95

Db 582 GNAGMMLGLGHGATVSVILWVIAAALMLNRKLLDAPGALMTGVLAVAGTFLVFFDLV 641

Qy 96 ALDGITKKAFFOWTGWKGVMLAGVVTMAVTRUTE 129

Db 642 ALSGVPRAIAPFLS--GIALLTIAAMRGRTSE 672

RESULT 13

US-09-815-242-10165

; Sequence 10165, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:13:37 ; Search time 33.8253 Seconds
(without alignments)
7815.311 Million cell updates/sec

Title: US-09-508-487-20
Perfect score: 862
Sequence: 1 gattttcattgatccag.....atatattttctctgtaa 862

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	49.2	5.7	6243	2	US-09-056-075-1
2	49	5.7	1422	1	US-08-319-704-5
3	49	5.7	3077	3	US-09-410-028-3
4	47.6	5.5	6243	2	US-09-056-075-1
5	47.2	5.5	2251	4	US-08-991-677-11
6	46.2	5.4	2960	3	US-08-913-842-3
7	45.8	5.3	734	3	US-09-014-583-1
8	45.8	5.3	1667	1	US-08-485-284A-1
9	45.2	5.2	782	4	US-08-998-416-224
C 10	45.2	5.2	3618	1	US-07-872-678A-36
C 11	44.2	5.1	20674	4	US-09-641-638-651
12	44	5.1	837	4	US-08-998-416-288
13	44	5.1	834	4	US-08-998-416-534
C 14	44	5.1	1850	3	US-08-617-860B-32
C 15	44	5.1	4098	2	US-08-605-106-4
16	44	5.1	5064	4	US-08-936-107A-8
C 17	43.8	5.1	1972	1	US-08-463-048-1
C 18	43.8	5.1	1972	1	US-08-463-229-1
C 19	43.8	5.1	19124	2	US-08-302-891-1
C 20	43.8	5.1	19124	2	US-08-487-826B-13
21	43.6	5.1	688	4	US-08-998-416-972
22	43.6	5.1	4818	3	US-08-817-926-27
23	43.4	5.0	246240	2	US-08-724-394A-20
24	43.4	5.0	246240	2	US-08-724-394A-21
25	43.4	5.0	246240	2	US-08-724-394A-22
C 26	42.8	5.0	98844	4	US-09-791-211-10
27	42.4	4.9	658	4	US-08-998-416-595

C 28	42.4	4.9	740	4	US-08-998-416-971	Sequence 971, Appl
29	42.4	4.9	3138	1	US-07-867-106-4	Sequence 4, Appli
30	42.4	4.9	5852	1	US-07-867-106-2	Sequence 2, Appli
C 31	42	4.9	3395	4	US-09-103-478-3	Sequence 3, Appli
C 32	42	4.9	3395	4	US-09-103-478-3	Sequence 3, Appli
C 33	42	4.9	7560	4	US-09-103-478-4	Sequence 4, Appli
C 34	42	4.9	7560	4	US-09-103-478-4	Sequence 4, Appli
35	41.8	4.8	636	4	US-08-998-416-1137	Sequence 1137, Ap
36	41.8	4.8	1850	3	US-08-617-860B-32	Sequence 32, Appli
37	41.8	4.8	4098	2	US-08-605-106-4	Sequence 4, Appli
C 38	41.6	4.8	1826	4	US-09-286-691-11	Sequence 11, Appl
C 39	41.6	4.8	1826	4	US-09-687-147-11	Sequence 11, Appl
40	41.6	4.8	3618	1	US-07-872-678A-36	Sequence 36, Appl
41	41.4	4.8	8920	2	US-08-446-855A-1	Sequence 1, Appli
42	41.4	4.8	8920	4	US-09-150-741-1	Sequence 1, Appli
C 43	41.4	4.8	11459	4	US-09-462-136-7	Sequence 7, Appli
44	41	4.8	663	4	US-08-998-416-191	Sequence 191, App
45	41	4.8	860	4	US-08-998-416-287	Sequence 287, App

ALIGNMENTS

RESULT 1
US-09-056-075-1/c
; Sequence 1, Application US/09056075
; Patent No. 5955368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite
; TITLE OF INVENTION: Expression System for Clostridium
; TITLE OF INVENTION: Species
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/056,075
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95238
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3770..4013
; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orit) from
; OTHER INFORMATION: plasmid RP4"
; US-09-056-075-1

Query Match 5.7%; Score 49.2; DB 2; Length 6243;
Best Local Similarity 58.0%; Pred. No. 0.046;

	Query Match	5.7%	Score 49;	DB 1;	Length 1422;
	Best Local Similarity	46.8%;	Pred. No. 0.038;		
	Matches 154;	Conservative 0;	Mismatches 175;	Indels 0;	Gaps 0;
56	AAAAGAATGAAATTCGTAATTTGGTCGCAAAATCAAGAGCTCTATTGGGAAGCGAATT	115			
91	ATAATTAATTTATGGATAAATTTATAGCATAGATTATATCTCTATATAGAAAAATAT	150			
116	TCAAGTAATACTTTGAAAAAGGTTAAATTTAAATAGCTTTTAAAAACCTTTTAAATTTTC	175			
151	TAAAAAATAATTTTATAAATCTTTTAAATTTTAAATTTTATTTATTTTCTGATGAA	210			
176	ATTAATATGTTACTATTAATACCAGTTTATAAAGAGGTTTTTATGAAATAAAATTTTAAT	235			

	Query Match	5.7%	Score 49	DB 3	Length 3077	
	Best Local Similarity	48.4%	Pred. No. 0.044			
	Matches 136	Conservative 0	Mismatches 145	Indels 0	Gaps 0	
QY	95	AGCPTATTGGGAGCGAAATTC	CAAGTAATACTTTG	AAAAAGTTAAATTTAAATAGTTT	154	
Db	1568	AGTTAAATGAGACTGACGAAC	TTTTATGGAAGTTAGTCTTAAGTTAGTTT	AGCTTGCTGTA 1627		
QY	155	TAAAAACCTTTTTPAAATTC	ATTAATGATTTACTATAATAC	CGAGTTTATAAAGAGGT	214	
Db	1628	GAGAAGCTCAATTTATTTT	ACTTTCTTATTTTCATCTTA	TAGTGTCTATGAAAAAGT	1687	
QY	215	TTTTATCAATAAAATTTT	TAAATGTGTTTTTGCTAG	CCCTTTTGTTTTTCTAGCTTGC	274	
Db	1688	TGATCTAAACAGGACTTAA	AAATGTTTTTAACTACCTT	CTCTCTATGGTGTGTTTTTA 1747		
QY	275	TCAAGCTGATGATTCTTAA	AGCGCTTTTAAATTTGGG	AGCGGAGAAAAACCTTTT	AGCTTA 334	
Db	1748	TGAACCTTAAAGTTATAC	TACATTCGCTAAATGAT	ACAAAGCTAGTGGAAAGT	TGGGAATT 1807	
QY	335	TGAACTAGTAAGAAAGAT	CTCTATTGTG	CGCAATTTTTATTGA 375		

Db 1808 AGAAATTTGAACAGTTATACAGTTACAGGGTGGTGGA 1848

RESULT 4

US-09-056-075-1

; Sequence 1, Application US/09056075

; Patent No. 5955368

; GENERAL INFORMATION:

; APPLICANT: Johnson, Eric A.

; APPLICANT: Bradshaw, Marite

; APPLICANT: Rood, Julian

; TITLE OF INVENTION: Expression System for Clostridium

; TITLE OF INVENTION: Species

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/056,075

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95238

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 608-251-5000

; TELEFAX: 608-251-9166

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6243 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 3770..4013

; OTHER INFORMATION: /note= "RP4 origin of DNA transfer (orig) from

; OTHER INFORMATION: plasmid RP4"

US-09-056-075-1

Query Match 5.5%; Score 47.6; DB 2; Length 6243;
Best Local Similarity 46.9%; Pred. No. 0.1;
Matches 183; Conservative 0; Mismatches 204; Indels 3; Gaps 1;

QY 19 AGAATTTGTAGATTTTCGACAATAAAGACATTAATAAGAAATTCGAATTCGTAATTT 78

Db 2733 AATATATATCATATATTTTCAATAATTAATACTATCTTAAACTCTTAAATAAAAAATAGAG 2792

QY 79 TATGGTCAAAATCAAGAGCTCTATTGGGAAGCGAATTTCAAGTAATACTTTGAAAAAAGT 138

Db 2793 TAAGTCCCAATTAACATTAATCTATTTTATGTTTTTAATTAATTAATTAATTAATAA 2852

QY 139 TAAATTAATAGTTTAAACCTTTTAAATTTTCAATTAATTAATTAATTAATTAATTAACCA 198

Db 2853 TATTTAACTAAATTAATGATTCCTTTTAAATTTTCAATTAATTAATTAATTAATTAATTA 2909

QY 199 GTTTTAATAAGAGTTTATTAATAAATTTTAAATTTTGTGTTGCTAGCCTTTTGTG 258

Db 2910 TATTACTATAATTAATTAACAAATTAATTTCTCATTTGTAATTAATTAATTAATTAATTA 2969

QY 259 TTTTCTTCTAGCTTTGCTCAAGCTGATGATTCCTAAAGCGCTTTTAAATTTGGAGCGGAG 318

Db 2970 ATTTTGTATTTTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3029

QY 319 AAAAATTTTGTAGCTTATCAAACTAGTAAGAAGATCCTTAATTTGTCCTATTTTATTGAACC 378

Db 3030 TATAAATTAATGATTAATTTTCCAAAGATATAAAGATTTTAATTTTCAATTTTAAACAATAC 3089

QY 379 TTTTGTAGGTTTGGGAATAGTTCTTTTGG 408

Db 3090 TTTTGTAAATTAATGTTTAAATTTAATTTG 3119

RESULT 5

US-08-991-677-11

; Sequence 11, Application US/08991677A

; Patent No. 6252135

; GENERAL INFORMATION:

; APPLICANT: Chiang, Vincent L

; APPLICANT: Carraway, Daniel T

; APPLICANT: Smeltzer, Richard H

; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms

; FILE REFERENCE: 50617

; CURRENT APPLICATION NUMBER: US/08/991.677A

; CURRENT FILING DATE: 1997-12-16

; EARLIER APPLICATION NUMBER: US 60/033,381

; EARLIER FILING DATE: 1996-12-16

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 2251

; TYPE: DNA

; ORGANISM: Pinus taeda

US-08-991-677-11

Query Match 5.5%; Score 47.2; DB 4; Length 2251;
Best Local Similarity 49.6%; Pred. No. 0.1;
Matches 121; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 50 ATATATAAAGAAATTTGAATTTGCTAAATTTTATGCTCAAAATCAAGAGCTCTATTGGGAAG 109

Db 619 ACTATAAATTTGAGTTGTGAATGAGTGTGTTTATGAGATGTTTAAATTTATATA 678

QY 110 CGAATTTCAAGTAATACTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTATA 169

Db 679 TGTAATTAATAATTTTATTTTGAATAACAATAATTAATTTGGATAAAAAATGTTTGT 738

QY 170 AATTTCAATTAATGTTTACTATAATACCAGTTTAAATAAGAGGTTTTTATGAATAAAT 229

Db 739 AATTTAGAGTAAATAATTTCAAAATCTAAAATAATTAACACACTATTATTTTAAAAAAT 798

QY 230 TTTAATTTGTTGTTTGTAGCTTTTGTGTTTTTCTAGCTTTTCTAGCTTTGCTCAAGCTGATGATC 289

Db 799 TGTGTTGTAATTTTATCTATATTAAGTTAAAAATTTAGAAAAAATTAATTTAAATTA 858

QY 290 TAAA 293

Db 859 TAAA 862

RESULT 6

US-08-913-842-3/C

; Sequence 3, Application US/08913842

; Patent No. 6028250

; GENERAL INFORMATION:

; APPLICANT: OHBA, Toshiharu

; APPLICANT: TAKAHASHI, Shuichi

; APPLICANT: ANMA, Yoshiko

; APPLICANT: ASADA, Kiyozo

; APPLICANT: KATO, Ikunoshin

; TITLE OF INVENTION: PLANT PROMOTER AND METHOD FOR GENE

; NUMBER OF SEQUENCES: 75

; CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 7th Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/913.842
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 07-073043
FILING DATE: 30-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/00777
FILING DATE: 26-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: OHBA-1
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2960 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-913-842-3

Query Match 5.4%; Score 46.2; DB 3; Length 2960;
Best Local Similarity 50.2%; Pred. No. 0.17;
Matches 114; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 30 AATTTTCGCAATAAAGACATATATAAAGAAATTCGAATTCCTAATTTTGGTCAAT 89
DB 1732 AATTTTATATATATAAATCTAAATCAATCAATATATATATTTATGCAAGAG 1673
QY 90 CAAGAAGCTCTATTGGGAAGCAATTCAGTAATACCTTGAAAAAGTTAAATTTAAAT 149
DB 1672 ATAGCAGGTAAATATCAAGCAGTTTAAATATATATTGATCAATGTAATTTGTTAA 1613
QY 150 AGTTTAAACCTTTTAAATTTCAATTAATGTTACTATAATACCAAGTTTAAATAAA 209
DB 1612 TTTCCTGTTATTCAAATTTCTTATTATAATGTTGCTAATGATATAAAGGTTTAAACTAA 1553
QY 210 GAGGTTTTATCAATAAATTTTAAATGTTGTTGCTAGCCCTTTTG 256
DB 1552 GGTGTTTTAACTAAATTTGGGATATTGGGTGTTCTAAATCCATTG 1506

RESULT 7
US-09-014-583-1
; Sequence 1, Application US/09014583
; Patent No. 6140077
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, Takeshi
; APPLICANT: SUZUKI, Tadashi
; APPLICANT: TOKUDA, Junko
; APPLICANT: KATO, No. 6140077uo
; APPLICANT: SAKAI, Yasuyoshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: TAKAHASHI, Hitoshi
; TITLE OF INVENTION: METHOD FOR PRODUCING PHYTASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014.583
FILING DATE: 28-JAN-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 200398/96
FILING DATE: 30-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 027066-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 734 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Candida boidinii
DEVELOPMENTAL STAGE: wild type
US-09-014-583-1

Query Match 5.3%; Score 45.8; DB 3; Length 734;
Best Local Similarity 49.8%; Pred. No. 0.16;
Matches 143; Conservative 0; Mismatches 142; Indels 2; Gaps 1;

QY 38 ACAATAAAGACATATTAAAGAAATTTGAAATTCCTAATTTTATGTCATAAATCAAGAAGC 97
DB 425 ATAAAAAATATATATTTAAATGAATTAATTCCTTTATTTAAATATATCGTTAATCT 484
QY 98 TCTATTGGGACGCAATTTCAAGTAATCTTTGAAAAAGTTAAATTTAAATAGTTTAA 157
DB 485 TTTAAATTCCTATTATTTTAAATTCCTTTTATCATAGTTATCATATATCAATAT 542
QY 158 AAACCTTTTAAATTCATTAATATGTTACTATATATACAGTTTAAATAAAGAGGTTT 217
DB 543 TAACATAGATACACAATTTATTTTATTTATTCATATTTTAAATATTTTAAATATTT 602
QY 218 TATGAATAAATTTTAAATGTTGTTGTTGCTAGCCTTTTGTGTTTCTAGCTTCTCA 277
DB 603 TTTAAATATATCTTTAAATTAATTTTACGAATATACAAATTTTAAACGACTTTCTT 662
QY 278 AGCTGATGATCTTAAAGCGCTTTTAAATTTGGGAGGGGAGAAAC 324
DB 663 TTTTACGAATTTTACGAACCTTTTAAAAAACAACAAAAAACAAC 709

RESULT 8
US-08-485-284A-1
; Sequence 1, Application US/08485284A
; Patent No. 5750372
; GENERAL INFORMATION:
; APPLICANT: SAKAI, YASUYOSHI
; APPLICANT: TANI, YOSHIKI
; APPLICANT: SHIBANO, YUJI
; APPLICANT: KONDO, HIROTO
; APPLICANT: HATANAKA, HARUYO
; TITLE OF INVENTION: NOVEL VECTOR HAVING PROMOTER THAT IS

TITLE OF INVENTION: INDUCIBLE BY METHANOL AND/OR GLYCEROL
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D. C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,284A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 43361/1992
FILING DATE: 28-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/025,416
FILING DATE: 01-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: WHITE JR., PAUL E.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 217755/FPS38209US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 1:
LENGTH: 1667 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: genomic DNA
US-08-485-284A-1

Query Match 5.3%; Score 45.8; DB 1; Length 1667;
Best Local Similarity 49.8%; Pred. No. 0.19;
Matches 143; Conservative 0; Mismatches 142; Indels 2; Gaps 1;
Qy 38 ACAATAAGACATTTAAAGAATTCGAATTCGCTAATTTTATGCTCAATCAAGAAGC 97
Db 1366 ATAAAAAATTATATATTTAAATGAATTAATTCCTTTATTTTAAATAATATCGTTAATCT 1425
Qy 98 TCTATTGGGAAGCGAATTTCAAGTAATACCTTTGAAAAAAGTTAAATTTAAATAGTTTAA 157
Db 1426 TTTAAATTCATTTTATTTTAAATTCCTTTTATATCATAGTTATCATATAACAA--TTATA 1483
Qy 158 AACCTTTTAAATTCATTAATATGTTACTATAATACCAAGTTTAAATAAGAGGTTT 217
Db 1484 TAACATAGATACACAATTTATTTATTTATCATATATTTTAAAAATATGATTATTT 1543
Qy 218 TATGAATAAATTTTAAATTTGTTTTCGTAGCCCTTTTGTGTTTTCCTAGCTTTGCCTCA 277
Db 1544 TTAATAATATCTCTAAATTAATTAATTTTACGAATATACAAATTTTAAACGACTTTCTTT 1603
Qy 278 AGCTGATGATCTCAAAAGCGCTTTTAAATTTGGGCGGAGAAAAAC 324
Db 1604 TTTTAAACGAATTTTAAACGAATTTTAAAAAACAACAAAAAACAAC 1650

RESULT 9
US-08-998-416-224/c
; Sequence 224, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer

APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPH
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PE/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 224:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAGL208UP
US-08-998-416-224

Query Match 5.2%; Score 45.2; DB 4; Length 782;
Best Local Similarity 48.1%; Pred. No. 0.22;
Matches 128; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
Qy 40 AAATAAGACATTTAAAGAATTCGAATTCGCTAATTTATGGTCAATCAAGAGCTC 99
Db 464 AAATAAGATATTTACTAGACTATAAATAAATAATTTATGAAGGAACATTTAATAATTA 405
Qy 100 TATTGGGAAGCGAATTTCAAGTAATACCTTTGAAAAAAGTTAAATTTAAATAGTTTAA 159
Db 404 TATATATCAGGGGAAGTAATTAATAATTTATATATAAATTAATATATATATATATAT 345
Qy 160 ACCCTTTTAAATTCATTAATATGTTACTATAATACCAAGTTTAAATAAGAGGTTT 219
Db 344 TTAATAAACTTAATAATCTATTTTATTAATAAATGGTATATATTAATTAATAAAT 285
Qy 220 TGAATAAATTTTAAATTTGTTTGTGCTAGCCCTTTTGTGCTAGCTTTTGTGCTCAAG 279
Db 284 TAAATACATTTTAAATAAATATCTATATAGTAATATTTCTTATTTTATAACATTT 225
Qy 280 CTGATGATCTCAAAAGCGCTTTTAAAT 305
Db 224 ATGTTTATGTTAAATAGATAATAAT 199

RESULT 10
US-07-872-678A-36/c
; Sequence 36, Application US/07872678A

Patent No. 5541060
GENERAL INFORMATION:
APPLICANT: Bell, Graeme, et al.
TITLE OF INVENTION: DETECTION OF EARLY-ONSET
NON-INSULIN-DEPENDENT DIABETES MELLITUS
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: Post Office Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/872,678A
FILING DATE: 22-APRIL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coughlin, Daniel F.
REGISTRATION NUMBER: 36,111
REFERENCE/DOCKET NUMBER: ARCD016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 3618 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-872-678A-36

Query Match 5.2%; Score 45.2; DB 1; Length 3618;
Best Local Similarity 52.1%; Pred. No. 0.29;
Matches 101; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 46 AGACATTATTAAGAATGCTAATTTTATGGTCAATCAAGAGCTCTATTGG 105
DB 3030 AGATGTTAATGAAGTATTAATGCTTAACTATAAGACATAACAATAAGGATAAGTT 2971
106 GAAGCGAATTTCAAGTAATACCTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTT 165
DB 2970 AAAGTTAAATTTAAATTAATGTTAAATTAACCTTAAATTTTAAATATGTTAATAAT 2911
QY 166 TTTAAATTTCAATATGTTACTAATAATACCAAGTTTAAATAAGAGGTTTATGTAATA 225
DB 2910 TTTTAAATTTAAATTTAAATTAATAATAATATGTTAATTTAAATTAATGTTAATAAT 2851
QY 226 AATTTTAAATGTT 239
DB 2850 AAACCTTAAATTTT 2837

RESULT 11
US-09-641-638-651
Sequence 651, Application US/09641638
Patent No. 6432648
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueret, Lydia
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BLALELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GENSET.051CPI
CURRENT APPLICATION NUMBER: US/09/641,638

CURRENT FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 651
LENGTH: 20674
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1123..3123
OTHER INFORMATION: 5'regulatory region
NAME/KEY: exon
LOCATION: 3124..3297
OTHER INFORMATION: exon 1
NAME/KEY: exon
LOCATION: 3871..4072
OTHER INFORMATION: exon 2
NAME/KEY: exon
LOCATION: 5552..5633
OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 5758..5880
OTHER INFORMATION: exon 4
NAME/KEY: exon
LOCATION: 5996..6099
OTHER INFORMATION: exon 5
NAME/KEY: exon
LOCATION: 6349..6509
OTHER INFORMATION: exon 6
NAME/KEY: exon
LOCATION: 7379..7522
OTHER INFORMATION: exon 7
NAME/KEY: exon
LOCATION: 8645..8854
OTHER INFORMATION: exon 8
NAME/KEY: exon
LOCATION: 12254..12340
OTHER INFORMATION: exon 9
NAME/KEY: exon
LOCATION: 12854..13023
OTHER INFORMATION: exon 10
NAME/KEY: exon
LOCATION: 13308..13429
OTHER INFORMATION: exon 11
NAME/KEY: exon
LOCATION: 16567..16667
OTHER INFORMATION: exon 12
NAME/KEY: exon
LOCATION: 16775..16945
OTHER INFORMATION: exon 13
NAME/KEY: exon
LOCATION: 17063..17554
OTHER INFORMATION: exon 14
NAME/KEY: misc_feature
LOCATION: 17555..20674
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 1128
OTHER INFORMATION: 10-508-191 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1182
OTHER INFORMATION: 10-508-245 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1559
OTHER INFORMATION: 10-509-284 : polymorphic base C or T

;; APPLICANT: Mohr, Christine
;; APPLICANT: Wendland, Jurgén
;; APPLICANT: Knechtle, Philipp
;; APPLICANT: Reibischung, Corinne
;; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
;; TITLE OF INVENTION: AND USES THEREOF
;; NUMBER OF SEQUENCES: 1152
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 6239264artis Corporation
;; STREET: 3054 Cornwallis Road
;; CITY: Research Triangle Park
;; STATE: No. 6239264th Carolina
;; COUNTRY: USA
;; ZIP: 27709
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/998,416
;; FILING DATE: 24-DEC-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: CH 0016/97
;; FILING DATE: 31-DEC-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Meigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8587
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 288:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 837 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: PAG1241RP
US-08-998-416-288

Query Match 5.1%; Score 44; DB 4; Length 837;
Best Local Similarity 46.0%; Pred. No. 0.4;
Matches 184; Conservative 0; Mismatches 215; Indels 1; Gaps 1;
QY 24 TTGTAGAATTTTCGCAATAAAGACATTAATAAAGAAATTCGAATTCGTAATTTATGG 83
DB 165 TTATTAATAATTTATATAATATAAATAATATTTAATATGAATCTATTTAGTCTATGT 224
QY 84 TCAAAATCAAGAGCTCTATTGGGAAGCGAATTTCAAGTAATACTTTGAAAAAGTTAAAT 143
DB 225 TCAATTTTAAATTTAGTTATTAATAATATTTAGATATTTATTTCTTTAATAAATA 284
QY 144 TTAATATAG-TTTTAAAAACCTTTTAAATTTCAATTAATGATGCTACTAATACCAGTTT 202
DB 285 TTAATATAGATATCAATAATAATATATTTATTTAATTTGTTTATTAATAAATAATAT 344
QY 203 TAATAAGAGGTTTATTAAGCAATAAATTTTAAATTTGTTGTTTCGAGCCTTTGTGTTT 262
DB 345 TTATTTATTAAGAAGATTTAAATTTATTTAAATATTGTAATTTATTTATTTATTAATAT 404
QY 263 TTCTAGCTTTGCTCAAGCTGATGATCTTAAAGCCCTTTTAAATTTGGGAGCGGAGAAA 322
DB 405 CTATTTTATAAATATATGTTGATTTATTTATTTAATCTTTTATAGAATTTATTTAT 464
QY 323 ACTTTTATGATGAACCTAGTAAGAAAGACCTATTTGTCATTTTATTTGAACTTTT 382
DB 465 AAAATTAATTTAACTTTAAATTTCTTATTTATTAATTTTATATTTATTAATAAATATAT 524
QY 383 TTAGGGTTGGAATAGGTTCTTTTGCTCAAGGATATTT 422

DB 525 TTCATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 564
RESULT 13
US-08-998-416-534
; Sequence 534, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgén
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 534:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1372UP
US-08-998-416-534

Query Match 5.1%; Score 44; DB 4; Length 854;
Best Local Similarity 50.4%; Pred. No. 0.4;
Matches 134; Conservative 0; Mismatches 130; Indels 2; Gaps 1;
QY 40 AAATAAGACATTTATAAAGAAATTCGAATTCGTAATTTATGGTCAAAATCAAGAGCTC 99
DB 303 AAATAAGAAATTTACTAGACTATAAATAAATAATATTTATGAAAGGTACATTTAATAATTA 362
QY 100 TATTCGGAGCCAAATTCAGTAATACTTTGAAAAAGTTAAATTTAAATAGCTTTTAAAA 159
DB 363 TATATATCAATCAAGTAATAATTTAATAATTTATTAATAATTTATATATCTACTTAATAT 422
QY 160 ACCTTTTAAATTTTCAATTAATGTTACTATAATACCAGTTTAAATAAGAGGTTTATTA 219
DB 423 TAATAAATCTAAT--AATCTATTTATTAATAAATAAGTATATTTAAATTTAAATAATAA 480
QY 220 TGAATAAATTTTAAATTTGTTGTTTCTAGCCCTTTTGCTAGCCCTTTTGCTAGCTTTC 279

Db	481	TTAAATACATATTAAATAAATATCTAAAGTAATTTCTTATTTATTTATACATTTAA	540
QY	280	CATGATGATTTCTAAAGCGCTTTTAAAT	305
Db	541	ATGTTTTTATGTTAAATAGATAATAAT	566

RESULT 14	
US-08-617-860B-32/c	
; Sequence 32, Application US/08617860B	
; Patent No. 6133506	
; GENERAL INFORMATION:	
; APPLICANT: Typfer, R., Bautor, J., Bothmann, H., Filsak, E.,	
; APPLICANT: Hvrice-Grandpierre, C., Klein, B., Martini, N.,	
; APPLICANT: Mller, A., Schulte, W., Voetz, M., Walek, J.,	
; APPLICANT: Schell, J.	
; TITLE OF INVENTION: Promoters	
NUMBER OF SEQUENCES: 35	
QY 92 AGAAGCTCTATTGGGAAGCGAATTTCAAGATAACTATCTTGAAAAAGTTTAAATTTAAATAG 151	
Db 354 AAAAAATATTTTTAAAAATTCACACAATATTTTAAAAATAAAAAAACTAATTTTAAAAATA - 296	
QY 152 TTTTAAAAACCCTTTTTTAAAATTCATTAATATGTTTACTATATATACAGGTTTAAATFAAGA 211	
Db 295 --TTTAAAAAATTTTAAAAAACFAATTTTATATAATTTTAAAAATAATTTTTTTTAAAAAGAAT 238	
QY 212 GCTTTTATGAATAAATTTTAAATTTGTTGTTT 243	
Db 237 TTAAAAATATATTTTAAATTTTAAAAAATATTTT 206	

RESULT 15
US-08-605-106-4/c
; Sequence 4, Application US/08605106
; Patent No. 5910631
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.

```
NAME/KEY: CDS
LOCATION: join(1797..2294, 2658..2791, 2898..3011, 3132
FEATURE: ..3303, 3391..3459, 3672..3941)
NAME/KEY: Startcodon
LOCATION: 1797..1799
FEATURE:
NAME/KEY: exon II
LOCATION: 1787..2294
FEATURE:
NAME/KEY: intron II
LOCATION: 2295..2657
FEATURE:
NAME/KEY: exon III
LOCATION: 2658..2791
FEATURE:
NAME/KEY: intron III
LOCATION: 2792..2897
FEATURE:
NAME/KEY: exon IV
LOCATION: 2898..3011
FEATURE:
NAME/KEY: intron IV
LOCATION: 3012..3131
FEATURE:
NAME/KEY: exon V
LOCATION: 3132..3303
FEATURE:
NAME/KEY: intron V
LOCATION: 3304..3390
FEATURE:
NAME/KEY: exon VI
LOCATION: 3391..3459
FEATURE:
NAME/KEY: intron VI
LOCATION: 3460..3671
FEATURE:
NAME/KEY: exon VII
LOCATION: 3672..3941
FEATURE:
NAME/KEY: Stopcodon
LOCATION: 3942..3944
US-08-605-106-4
```

```
Query Match 5.1%; Score 44; DB 2; Length 4098;
Best Local Similarity 53.8%; Pred. No. 0.54;
Matches 114; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

QY 32 TTTCGACAAATAAGACATTTATTAAGAAGATTGAAATTCGTAATTTTATGTCAAATCA 91
Db 414 TTTCGAATATTCAAAAATTCAAAAATATTTAAAACTTTTAAATATATTTTAAACC 355
QY 92 AGAAGCTCTATGGGAGCGAATTTCAAGTATACCTTTCGAAAAAGCTTAAATTTAAATAG 151
Db 354 AAAAAATATTTTAAAAATTCACAAATATTTTAAAAATAAAAAAACTAATTTTAAAAATA- 296
QY 152 TTTTAAAAACCTTTTAAATTCATTAATATGTTACTATATACCAAGTTTAAATAAGA 211
Db 295 --TTTAAAAATTTTAAAAACCTAATTTTATATATATTTTAAAAATATTTTAAAAAGAAT 238
QY 212 GGTTTTATGAATAAATTTTAAATGTTGTTT 243
Db 237 TTAATAATATTTAAATTTTAAAAAATATTTT 206
```

Search completed: March 15, 2003, 07:44:04
Job time : 71.8253 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:15:46 ; Search time 1448.7 seconds
(without alignments)
12999.061 Million cell updates/sec

Title: US-09-508-487-22
Perfect score: 749
Sequence: 1 tagaattttcaacaaataaaa.....atttatctagaaaaatgggtg 749

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 24791104 seqs, 12571243825 residues 49582208
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	749	100.0	749	15	US-09-153-447-22	Sequence 22, Appl
2	749	100.0	749	15	US-09-156-447-22	Sequence 22, Appl
3	749	100.0	749	19	US-09-508-487-22	Sequence 22, Appl
4	607	81.0	862	15	US-09-153-447-20	Sequence 20, Appl
5	607	81.0	862	15	US-09-156-447-20	Sequence 20, Appl
6	558.2	74.5	759	15	US-09-153-447-18	Sequence 18, Appl
7	558.2	74.5	759	15	US-09-156-447-18	Sequence 18, Appl
8	558.2	74.5	759	19	US-09-508-487-18	Sequence 18, Appl
9	98.4	13.1	1761	1	PCT-US98-12764-63	Sequence 63, Appl
c 10	98.4	13.1	1761	32	US-09-830-228-63	Sequence 63, Appl
c 11	91.6	12.2	9399	1	PCT-US98-12764-14	Sequence 14, Appl
12	91.6	12.2	9399	32	US-09-830-228-14	Sequence 14, Appl
13	82.2	11.0	7074	1	PCT-US98-12764-19	Sequence 19, Appl
14	82.2	11.0	7074	32	US-09-830-228-19	Sequence 19, Appl
15	66	8.8	7110	42	US-10-221-613-251	Sequence 251, Appl
16	61.4	8.2	417	17	US-09-362-510-57521	Sequence 57521, A
c 17	61.4	8.2	417	17	US-09-362-510-57521	Sequence 57521, A
c 18	61.4	8.2	417	34	US-09-904-013-57521	Sequence 57521, A
19	59.4	7.9	402	33	US-09-873-402A-57873	Sequence 57873, A
c 20	59.4	7.9	402	33	US-09-873-402A-57873	Sequence 57873, A
c 21	59.4	7.9	402	64	US-60-209-830-49195	Sequence 49195, A

Sequence 2899, Ap
Sequence 267508,
Sequence 266370,
Sequence 22020, A
Sequence 71765, A
Sequence 71765, A
Sequence 267508,
Sequence 65625, A
Sequence 422565,
Sequence 422565,
Sequence 1501, Ap
Sequence 2437, Ap
Sequence 2437, Ap
Sequence 2437, Ap
Sequence 53, Appl
Sequence 681, Appl
Sequence 389, Appl
Sequence 74654, A
Sequence 73521, A
Sequence 190, Appl
Sequence 51, Appl
Sequence 5933, Ap
Sequence 5933, Ap
Sequence 40093, A

ALIGNMENTS

US-09-153-447-22
; Sequence 22, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; EARLIER FILING DATE: 1998-09-15
; EARLIER FILING DATE: 1997-09-10
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Borrelia garinii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (192)..(725)
US-09-153-447-22
Query Match 100.0%; Score 749; DB 15; Length 749;
Best Local Similarity 100.0%; Pred. No. 4.4e-122;
Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTTGAAATTTGCTAATTTTATGTTA 60
DB 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTTGAAATTTGCTAATTTTATGTTA 60
QY 61 AATCAAGAAGCTCTATTGGTAAAGCAATTTTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
DB 61 AATCAAGAAGCTCTATTGGTAAAGCAATTTTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
QY 121 AATAGTTCACAAAGCCCTTTTAAATTTTCAATTAATGCTACCATTAATACCAGTTTAAATA 180
DB 121 AATAGTTCACAAAGCCCTTTTAAATTTTCAATTAATGCTACCATTAATACCAGTTTAAATA 180
QY 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGTATCTTTTGTGCTTTTCTTA 240
DB 181 AAGGGGTTTTATGAATAAGTTTTTAAATTTTATTTTGTATCTTTTGTGCTTTTCTTA 240

QY 241 GTTTTGTCTCAAGATGATCTTAAAGACACTTTTAAATCGGAGCGGAGAGAAAAATTTTGG 300
DB 241 GTTTTGTCTCAAGATGATCTTAAAGACACTTTTAAATCGGAGCGGAGAGAAAAATTTTGG 300
QY 301 TTTATGAACATAATAAGAAATCTCTTGTACCATTTTATGAACCTTTTATTTAGGTT 360
DB 301 TTTATGAACATAATAAGAAATCTCTTGTACCATTTTATGAACCTTTTATTTAGGTT 360
QY 361 TCGGGATAGTCTCTTTTGTCTCAAGGAGATATCTTGGAGGTTCTCTTATTTCTTGGATTG 420
DB 361 TCGGGATAGTCTCTTTTGTCTCAAGGAGATATCTTGGAGGTTCTCTTATTTCTTGGATTG 420
QY 421 ATCGGTTTGTATAGGTTTAAATCTTACAGGAGCTTATTTGGACATCAAGGATTTTGATA 480
DB 421 ATCGGTTTGTATAGGTTTAAATCTTACAGGAGCTTATTTGGACATCAAGGATTTTGATA 480
QY 481 ATATGCTTAAAGAAAGCTGATTTTAAAGTGGACTTGGGTAAGGAATGATTTGGCAGGTG 540
DB 481 ATATGCTTAAAGAAAGCTGATTTTAAAGTGGACTTGGGTAAGGAATGATTTGGCAGGTG 540
QY 541 TGCTTACTATGGCTGTGACAAAGATTTGACAGAAATTTGTTCTTCCATTTACATTTTCTAATA 600
DB 541 TGCTTACTATGGCTGTGACAAAGATTTGACAGAAATTTGTTCTTCCATTTACATTTTCTAATA 600
QY 601 ATTATACAGGAAGCTGAAATATAGTCTTAAATATAGCTTTGGAGGATTTGAGCCTAGTT 660
DB 601 ATTATACAGGAAGCTGAAATATAGTCTTAAATATAGCTTTGGAGGATTTGAGCCTAGTT 660
QY 661 TTGATATTACATGGGCAAGCTAGTGTCTTGTGTTTGGAGTGTCTTTTCAAAAAAGCT 720
DB 661 TTGATATTACATGGGCAAGCTAGTGTCTTGTGTTTGGAGTGTCTTTTCAAAAAAGCT 720
QY 721 ATTAATTTTATTTATCTAGAAATGGGTG 749
DB 721 ATTAATTTTATTTATCTAGAAATGGGTG 749

RESULT 2

US-09-156-447-22
; Sequence 22, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Borrelia garinii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (192)..(725)
US-09-156-447-22

Query Match 100.0%; Score 749; DB 15; Length 749;
Best Local Similarity 100.0%; Pred. No. 4.4e-122;
Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTTGAAATTTGCTAATTTTATGTTA 60
DB 1 TAGAATTTTCAACAAATAAGATATTGTTAAAGAAATTTGAAATTTGCTAATTTTATGTTA 60
QY 61 AATCAAGAAGCTCTATTGGTAAAGCAATTTTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
DB 61 AATCAAGAAGCTCTATTGGTAAAGCAATTTTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 22:48:52 ; Search time 48.9852 Seconds
(without alignments)
10731.744 Million cell updates/sec

Title: US-09-508-487-22
Perfect score: 749
Sequence: 1 tagaatttcaacaataaa.....atttatctagaaatgggtg 749

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues
Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	52	6.9	416	10	US-09-960-352-4584
2	48.4	6.5	416	10	Sequence 4584, Ap
3	48.2	6.4	2000	9	US-09-938-842A-5004
4	48	6.4	1044	9	Sequence 5004, Ap
5	48	6.4	1051	9	Sequence 10, Appl
6	48	6.4	1140	9	Sequence 22, Appl
C 7	47.2	6.3	640681	10	US-09-998-059-29
8	46.8	6.2	2251	10	US-09-790-988-1
9	45.8	6.1	385	10	US-09-796-256A-11
10	45.8	6.1	419	10	US-09-960-352-1739
C 11	45.2	6.0	3272	10	US-09-960-352-11234
12	44.2	5.9	376	10	US-09-796-348-18
13	44.2	5.9	640681	10	US-09-960-352-5087
C 14	44	5.9	513509	9	US-09-790-988-1
15	43.8	5.8	516	10	US-09-754-853A-4
16	43.4	5.8	397	10	US-09-960-352-5785
C 17	43.4	5.8	411	10	US-09-960-352-13784
18	43.4	5.8	413	10	Sequence 13784, A
19	43.4	5.8	431	10	Sequence 14521, A
					Sequence 2919, Ap
					Sequence 5558, Ap

C 20	43.4	5.8	335913	9	US-09-754-853A-2	Sequence 2, Appl
C 21	43.4	5.8	335913	9	US-09-754-853A-3	Sequence 3, Appl
C 22	43	5.7	479	10	US-09-960-352-12872	Sequence 12872, A
C 23	43	5.7	1367	9	US-09-938-842A-4959	Sequence 4959, Ap
C 24	43	5.7	4161	9	US-10-098-841-135	Sequence 135, App
25	42.8	5.7	446	10	US-09-960-352-3400	Sequence 3400, Ap
C 26	42.8	5.7	335913	9	US-09-754-853A-2	Sequence 2, Appl
C 27	42.8	5.7	335913	9	US-09-754-853A-3	Sequence 3, Appl
C 28	42.6	5.7	449	10	US-09-960-352-2412	Sequence 2412, Ap
C 29	42.6	5.7	73308	10	US-09-954-456-2276	Sequence 2276, Ap
C 30	42.4	5.7	335	10	US-09-960-352-13489	Sequence 13489, A
C 31	42.4	5.7	203654	10	US-09-820-905-3	Sequence 3, Appl
C 32	42	5.6	2000	9	US-09-938-842A-3655	Sequence 3655, Ap
C 33	42	5.6	2000	9	US-09-938-842A-5073	Sequence 5073, Ap
C 34	41.8	5.6	11474	10	US-09-816-028A-1	Sequence 1, Appl
C 35	41.6	5.6	291	10	US-09-960-352-1243	Sequence 1243, Ap
C 36	41.6	5.6	317	10	US-09-960-352-2699	Sequence 2699, Ap
C 37	41.6	5.6	368	10	US-09-834-975-47	Sequence 47, Appl
C 38	41.6	5.6	750	10	US-09-770-149-9	Sequence 9, Ap
C 39	41.6	5.6	5046	10	US-09-725-735A-13	Sequence 13, A
C 40	41.4	5.5	3303	10	US-09-796-348-19	Sequence 19, Appl
C 41	41.4	5.5	3315	10	US-09-796-348-17	Sequence 17, Appl
C 42	41.4	5.5	513509	9	US-09-754-853A-4	Sequence 4, Appl
C 43	41.2	5.5	2251	10	US-09-796-256A-11	Sequence 11, Appl
C 44	41	5.5	347	10	US-09-960-352-3371	Sequence 3371, Ap
C 45	41	5.5	1058	10	US-09-804-682-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1

US-09-960-352-4584/C
; Sequence 4584, Application US/09960352
; Patent NO. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4584
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11
US-09-960-352-4584

Query Match 6.9%; Score 52; DB 10; Length 416;

Best Local Similarity 51.8%; Pred. No. 0.29;

Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Qy	95	TAAACAATTTGAAAAAGTTAAATTTAAATAGTTCACAAAGCCTTTTAAATTTTCATTAA	154
Db	288	TTAAATTTTAAAAATTTTCTTTTAAAAATTTTAAAAATTTTAAAAATTTT	229
Qy	155	TATGCTACCAATAACACAGTTTAAATAAGGGTTTTTATGATAGTTTAAATTTTAT	214
Db	228	TTTTTATGAATTTTAAATTTTATATATTTTAAATTTTAAAAATCTTTTTTTTTTT	169
Qy	215	TTTGGTATCTTTTGTGCTTTTCTAGTTTTCCTCAAGATGATCTAAAAGCATTAA	274
Db	168	TTTTTTTTTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTCT	109
Qy	275	TCTGGGACGGGAGAAAAATTTTGGTTTATGAACACTAATAAGAAAGA	322
Db	108	ATTTTTCATTAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATA	61

RESULT 2
US-09-960-352-4584
; Sequence 4584, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4584
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-016-Q1-K1-E11
US-09-960-352-4584

Query Match 6.5%; Score 48.4; DB 10; Length 416;
Best Local Similarity 51.4%; Pred. No. 1.3;
Matches 112; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 11 AACAAATAAGATATGTTTAAAGAAATGAAATTCGTAATTTATGTTAAATCAAGAAG 70
DB 146 AATAAATAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAGATTTTAAAAATTAATAT 205
QY 71 CTCATTGTAAGCAATTCGAGTAAACAATTTGAAAAAGTTTAAATTTAAATAGTTCCA 130
DB 206 AAAAAATTAATAATTCATAAAAAAAAAAATAATTTAAAAAATAATTTTAAATTTTAA 265
QY 131 AAAGCCTTTTAAATTCATTAATGCTACCATATACCAGTTTAAATAAGGGGTTTT 190
DB 266 AGAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 325
QY 191 TATGAATAAGTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 228
DB 326 TAAAAATAAATTTTAAATAATATATGCGTAATTTAT 363

RESULT 3
US-09-938-842A-5004
; Sequence 5004, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 5004
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-5004

Query Match 6.4%; Score 48.2; DB 9; Length 2000;
Best Local Similarity 51.7%; Pred. No. 1.8;

Matches 134; Conservative 0; Mismatches 123; Indels 2; Gaps 1;
QY 2 AGAATTTTCAACAAATAAAGATATTCGTTAAAGAAATGAAATTCGCTAAATTTTATGGTTAA 61
DB 186 AGTACTATCATTTAGTAAATCTCTCGAATATATAATATGATTCGCAAAAATATACATTA 245
QY 62 ATCAAGAAGCTCTATTGGTAAGCAATTCGAGTAAACAATTTGAAAAAGTTTAAATTTAA 121
DB 246 ATTACCAATCTTTTCTCTATAATAAGTACCAGCTTAAATATACCATATAGCAGAAGTTATA 305
QY 122 ATAGTTTCCAAAAGCCCTTTTAAATTTCA--TTAATATGCTACCATATAATACCAGTTTAA 179
DB 306 TTCCTAAACAGATCGTTTCTTAAATTTCAATTTGTTTAAATTTATTTCAAGTACAAATCTTT 365
QY 180 AAAGGGGTTTTTATGAATAAGTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTCT 239
DB 366 GTAATGGTATTTTAAAAATGATTAATAACTTGATATTTTCATATAGTAGTTTGTGTT 425
QY 240 AGTTTGGCTCAAGATGATT 258
DB 426 TTGTAGGATCATGTAAT 444

RESULT 4
US-09-998-059-10
; Sequence 10, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1044
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-10

Query Match 6.4%; Score 48; DB 9; Length 1044;
Best Local Similarity 51.4%; Pred. No. 1.8;
Matches 111; Conservative 0; Mismatches 105; Indels 0; Gaps
QY 30 AAAAGAATTCAAAATTCGCTAAATTTTATGGTTAAATCAAGAAGCTCTATTGCTAAGCAATT 89
DB 504 ATATGAATGCTATTTTGTCTTAAACCTTAAATTTGTACTATTTGAAGGAATTTTCATTT 563
QY 90 TCGAGTAACAATTTGAAAAAGTTAAATTTAAATAGTTCAAAAGCCTTTTAAATTTTC 149
DB 564 TATTTTAACTGCTAGTATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 623
QY 150 ATTAATATGCTACCAATAATACCAGTTTAAATAAGGGGTTTTTATGAATAAGTTTAAAT 209
DB 624 TTAATTTCTTATTTATAAACAATAAAGTAATATTTAAATTTCTTTTAAATTAATAATAT 683
QY 210 TTTATTTTGGTATCTTTTGTGCTTTTCTAGTTTT 245
DB 684 TTTATTTTCAATAATTTATTTTGTCTTTTTTTTTTTT 719

RESULT 5
US-09-998-059-22
; Sequence 22, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:

```
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-22

Query Match          6.4%; Score 48; DB 9; Length 1051;
Best Local Similarity 51.4%; Pred. No. 1.8;
Matches 111; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Qy 30 AAAAGAATTGAAATGCTTAATTTTATGCTTAAATCAAGAAGCTCTATTGGTAAGCGAATT 89
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 504 ATATGAATGCTATTGCTCTTAAAACTTAAATTTGTAAGGAAATTTCAATTT 563
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 90 TCGAGTACAAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCCTTTTAAATTTTC 149
    || || || || || || || || || || || || || || || || || || || || ||
Db 564 TATTTTAAATGCTAGTAGATTTATATTAATTAATTAATTAATTAATTAATTAATTA 623
    || || || || || || || || || || || || || || || || || || || || ||

Qy 150 ATTAATGCTACCATATACAGTTTAAATAAAGGGTTTTTATGAATAAGTTTTTAATT 209
    || || || || || || || || || || || || || || || || || || || || ||
Db 624 TTAATCTTATTATTAACATAAGTAATTAATTTCTTAAATTAATAAATACATAT 683
    || || || || || || || || || || || || || || || || || || || || ||

Qy 210 TTTATTTGGTAATCTTTTGCTTTTCTTAGTTTT 245
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 684 TTTATTTTCATAAATTTATTTTGCTTTTTTTTTTTT 719
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-998-059-29
; Sequence 29, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:
; APPLICANT: Ohlrogge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-29

Query Match          6.4%; Score 48; DB 9; Length 1140;
Best Local Similarity 51.4%; Pred. No. 1.8;
Matches 111; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

Qy 30 AAAAGAATTGAAATGCTTAATTTTATGCTTAAATCAAGAAGCTCTATTGGTAAGCGAATT 89
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 ATATGAATGCTATTGCTCTTAAAACTTAAATTTGTAAGGAAATTTCAATTT 616
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 90 TCGAGTACAAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCCTTTTAAATTTTC 149
```

```
Db 617 TATTTTAAATGCTAGATTAGATTTTATAATTAATTAATTAATTAATGATGTAATAACTAAC 676
    || || || || || || || || || || || || || || || || || || || || ||
Qy 150 ATTAATATGCTACCATATACAGTTTAAATAAAGGGTTTTTATGAATAAGTTTTTAAAT 209
    || || || || || || || || || || || || || || || || || || || || ||
Db 677 TTAATTTCTTATTATTAACAACATAAAGTAAATTTTCTTAAATTAATAAATAACATAT 736
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 210 TTTATTTGGTAATCTTTTGCTTTTCTAGTTTT 245
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 737 TTTATTTTCATAAATTTATTTTGCTTTTTTTTTTTT 772
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match          6.3%; Score 47.2; DB 10; Length 640681;
Best Local Similarity 51.5%; Pred. No. 6;
Matches 134; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

Qy 87 ATTTGAGTACAAATTTGAAAAAGTTAAATTTAAATAGTTCGAAAGCCCTTTTAAAT 146
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546504 ATTTCAACAAAAATATAAAATTAATATATTTTATAAAAATAGTTCCTATTTTAA 546445
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 147 TTCATTAATATGCTACCATATACAGTTT---AATAAAGGGTTTTTATGAATAAGTTT 203
    || || || || || || || || || || || || || || || || || || || || ||
Db 546444 ATAAATTTTGTAGAAATATACATTTTAAACATCTATATATTTTTCGCGAAAT 546395
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 204 TTAATTTTATTTTGGTAATCTTTTGCTTTTCTAGTTTGTCTCAAGATGATCTAAA 263
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546384 GTAAATTTTAAATATTAATATTTGATCGATGATGATTTGATTTGATTTTATAT 546325
    ||||| ||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 264 AGCACTTTTAACTCGGAGCGGAGAAAAATTTTGGTTTATGAACATAAAGAAGAT 323
    || || || || || || || || || || || || || || || || || || || || ||
Db 546324 AGAAATAATTTCTCCTACAGAGCGGTATTTTGTTATAGGAATTAATCTCTGATGCC 546265
    || || || || || || || || || || || || || || || || || || || || ||

Qy 324 TCTCTGTGACCAATTTTATTT 343
    ||||| ||| ||| ||||| ||
Db 546264 ATACTTTTATATTTTATTT 546245
    ||||| ||| ||| ||||| ||

RESULT 8
US-09-796-256A-11
; Sequence 11, Application US/09796256A
; Patent No. US20020078477A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Vincent L
; APPLICANT: Carraway, Daniel T
; APPLICANT: Smeltzer, Richard H
; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
; FILE REFERENCE: 50617/c-3532.0
; CURRENT APPLICATION NUMBER: US/09/796,256A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US60/033381
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:25:06 ; Search time 23 Seconds
(without alignments)
1006.876 Million cell updates/sec

Title: US-09-508-487-21
Perfect score: 904
Sequence: 1 MNKFLIVLLAFCVSSFAQ.....DINMGQASALGFELSPKSY 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 598297 seqs, 130101922 residues

Total number of hits satisfying chosen parameters: 598297

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	89.5	9.9	450	6	US-10-282-122A-69822
2	87	9.6	505	5	US-09-134-000C-5340
3	87	9.6	505	5	US-09-134-000C-5340
4	87	9.6	505	6	US-10-282-122A-57091
5	83.5	9.2	442	6	US-10-369-493-12272
6	83.5	9.2	650	6	US-10-282-122A-73910
7	83	9.2	464	6	US-10-282-122A-71751
8	81.5	9.0	692	6	US-10-369-493-48
9	81.5	9.0	1291	6	US-10-354-774-42
10	81	9.0	305	5	US-09-988-067B-6
11	81	9.0	379	6	US-10-156-761-11883
12	81	9.0	479	6	US-10-282-122A-51378
13	79	8.7	486	6	US-10-366-683-19571
14	78.5	8.7	372	6	US-10-282-122A-78114
15	78.5	8.7	436	6	US-10-282-122A-66257
16	78.5	8.7	663	6	US-10-287-274-332
17	78.5	8.7	663	6	US-10-282-122A-56452
18	78.5	8.7	663	6	US-10-282-122A-75460
19	78.5	8.7	1291	6	US-10-354-774-40
20	78	8.6	540	6	US-10-156-761-13326
21	77	8.5	308	6	US-10-282-122A-53673
22	77	8.5	447	6	US-10-369-493-23305
23	77	8.5	792	6	US-10-156-761-9442
24	76.5	8.5	571	1	PCT-US02-32727-19854
25	76.5	8.5	571	5	US-09-978-825-19854
26	76.5	8.5	571	6	US-10-057-498-19854

27	76.5	8.5	652	6	US-10-282-122A-73132
28	76	8.4	269	5	US-09-950-084-6859
29	76	8.4	466	6	US-10-282-122A-44444
30	76	8.4	1226	6	US-10-369-493-1078
31	75.5	8.4	206	6	US-10-282-122A-59036
32	75.5	8.4	389	6	US-10-369-493-6417
33	75	8.3	375	5	US-10-369-493-11538
34	75	8.3	499	5	US-09-134-000C-6426
35	75	8.3	499	5	US-09-134-000C-6426
36	75	8.3	542	6	US-10-366-683-19270
37	74.5	8.2	214	6	US-10-282-122A-49274
38	74.5	8.2	429	6	US-10-282-122A-78316
39	74	8.2	203	6	US-10-282-122A-50724
40	74	8.2	403	6	US-10-156-761-12262
41	74	8.2	409	6	US-10-282-122A-59641
42	74	8.2	470	6	US-10-156-761-10352
43	74	8.2	479	6	US-10-282-122A-72971
44	74	8.2	561	6	US-10-282-122A-44555
45	73.5	8.1	337	6	US-10-282-122A-52786

ALIGNMENTS

RESULT 1

US-10-282-122A-69822
; Sequence 69822, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69822
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Pseudomonas syringae
US-10-282-122A-69822

Query Match 9.9%; Score 89.5; DB 6; Length 450;
Best Local Similarity 26.2%; Pred. No. 0.6;
Matches 49; Conservative 22; Mismatches 51; Indels 65; Gaps 11;
QY 5 LIVVLLAFC-VFSSFAQADSKSA-----FNL-----GAG-- 33
Db 49 MVIAPITCTVVGSIAGMNMKSVGTGYALLYFIVSTIALLIGLIVNVVQPGAGMN 108
QY 34 -----EKLAYETSKD-PIVPLNLFGLFGIGSFAQGDILGGLFGLGDAV-GIG 83
Db 109 IDVSTLDASKIAAYTAGDQSVGFILNVPTIVGAFANGDIL--QVLMFSEVIFGFA 165
QY 84 LILTGAY-----LD-----IKALDKNAPKAAFKWTWKGMMLAGAV-TMAVTRL 126
Db 166 LHLRSGYGRPVLDIFDRFAHVFMNIIMTKLAPIGAF----GAMAFITGAYGVSSLVQL 221
QY 127 TELIIPF 133
Db 222 GQLMICF 228
RESULT 2
US-09-134-000C-5340
; Sequence 5340, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5340
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5340

Query Match 9.6%; Score 87; DB 5; Length 505;
Best Local Similarity 23.3%; Pred. No. 1.2;
Matches 45; Conservative 26; Mismatches 68; Indels 54; Gaps 10;
QY 6 IYVLLAFCVFSSFAQA-----DDSKSAFNILGAGEKLLAYETSKKDPVLPFLN- 53
Db 85 IYVLLAFCVFSSFAQA-----DDSKSAFNILGAGEKLLAYETSKKDPVLPFLN- 53
QY 54 -----LFLG-----FGIGSFAQGDILGGLFGLGDAVIGLITGAYLDIKALDKNA 100
Db 142 IRTATVLIIGTATLAALIGAGG-----LGTFFILLGIDRNLSLIFIG-----ALSSAA 189
QY 101 PRAAFKW--TW-----GKGMMLAGAVTMAVTRLTEIIIPFTFANSYNRKLNLSIAFG- 152
Db 190 LAVLFNYGIHWLEKANGRRLIIGGTI-----LGLLLGGFFWNQOTSSKEKOLTIAAGKL 243
QY 153 GFEPSPDINMGOA 165
Db 244 GAEPDIIINMYKA 256
RESULT 3
US-09-134-000C-5340
; Sequence 5340, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5340
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5340
Query Match 9.6%; Score 87; DB 5; Length 505;
Best Local Similarity 23.3%; Pred. No. 1.2;
Matches 45; Conservative 26; Mismatches 68; Indels 54; Gaps 10;
QY 6 IYVLLAFCVFSSFAQA-----DDSKSAFNILGAGEKLLAYETSKKDPVLPFLN- 53
Db 85 IYVLLAFCVFSSFAQA-----DDSKSAFNILGAGEKLLAYETSKKDPVLPFLN- 53
QY 54 -----LFLG-----FGIGSFAQGDILGGLFGLGDAVIGLITGAYLDIKALDKNA 100
Db 142 IRTATVLIIGTATLAALIGAGG-----LGTFFILLGIDRNLSLIFIG-----ALSSAA 189
QY 101 PRAAFKW--TW-----GKGMMLAGAVTMAVTRLTEIIIPFTFANSYNRKLNLSIAFG- 152
Db 190 LAVLFNYGIHWLEKANGRRLIIGGTI-----LGLLLGGFFWNQOTSSKEKOLTIAAGKL 243
QY 153 GFEPSPDINMGOA 165
Db 244 GAEPDIIINMYKA 256
RESULT 4
US-10-282-122A-57091
; Sequence 57091, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1


```

; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 48
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
US-10-369-493-48

Query Match          9.0%; Score 81.5; DB 6; Length 692;
Best Local Similarity 24.6%; Pred. No. 6.7;
Matches 44; Conservative 17; Mismatches 49; Indels 69; Gaps

QY 39 YETSKDPIVPFL-----LNLFGLFGIGSFAQGDILGLFLILGFDVAGIGLILTGA 89
   || : ||| ||| | : | : | | | | | | | | | | | | | | | | | | | |
Db 84 YEDTYKKPIV-FLSPYGLHHTLLIYAG-GLG-FLAGDILKESDGLGFLIGVGFVWPG 140
   || : ||| ||| | : | : | | | | | | | | | | | | | | | | | | | |

QY 90 Y-----LPI-----KALDKNAKPAKF-----TWGK 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 141 YVKQRIVDGVQEDLDAQNQKELMPVKVLDKEG-----KWLKCYVVRDEKVFVGW-- 193
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 112 GMLAGAVTMVTRLTEIIPFTFANSYRKLNSLNIAFGGEPEDINMGASALGF 170
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 194 -----EVNVGRTKLYLDNTVNEPTWNRKISSRL-----YVPDKDLRLROOIVLGF 240
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
US-10-354-774-42
; Sequence 42, Application US/10354774
; GENERAL INFORMATION:
; APPLICANT: Williams, James A.
; Thalley, Bruce S.
; TITLE OF INVENTION: Multivalent Vaccine For Clostridium
; Botulinum Neurotoxin
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/354,774
; FILING DATE: 30-Jan-2003
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,159
; FILING DATE: 28-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: OPHD-02304
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1291 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-354-774-42

Query Match          9.0%; Score 81.5; DB 6; Length 1291;

```

```
Best Local Similarity 23.8%; Pred. No. 15;
Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;

Qy 47 IVPFLLNLFGLGSGFPAQGDILGGLILGFDV-----GIGLILGAYLD--- 92
||| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 620 IVPY-IGALNVG-NETAKGFENAFETAGASILLEFIPELLIPVVGAFLLSYIDNKNK 677
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 93 -IKALDKNAKPAKFKWTGKGMLAGAVTMAVTRL-----TEIIIPFT 134
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 678 IIKIDNALTRNEKWSMDYGLIQAQWLSVTNTQPTTIKEGMYKALNQQAALIEIIKYR 737
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 135 FANSYNRKLNLSIAFGFPFSPFDINMQA 165
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 738 Y-NIYSEKESKNINIDFINDSKNLNIGINQA 767
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :

RESULT 10
US-09-988-067B-6
; Sequence 6, Application US/09988067B
; GENERAL INFORMATION:
; APPLICANT: Haas, Rainer
; APPLICANT: Kleanthous, Harold
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Miller, Charles
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Odenbreit, Stefan
; APPLICANT: Meyer, Thomas
; TITLE OF INVENTION: Helicobacter Polypeptides and
; FILE OF INVENTION: Corresponding Polynucleotide Molecules
; FILE REFERENCE: 06132/040002
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US/09/988,067B
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-988-067B-6

Query Match 9.0%; Score 81; DB 5; Length 305;
Best Local Similarity 28.1%; Pred. No. 2.7;
Matches 57; Conservative 23; Mismatches 75; Indels 48; Gaps 14;

Qy 14 VFSFPAQADDSKAFN-LGAGEKLLAYETSKKDPVPPFLLNLFGLF---GIGSFAQGDIL 69
: | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 84 IASRFAGNGSGGLFNELSFYK---YFLGKK-RIIGFRHSILFFGYQLGGVGS-VPGSGL 138
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 70 GGFILGFD-----AVGIGLI---LTGAYLDIKALDKNAKPAF 105
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 139 IVLPYGFNTDILLINWTKRASOKYVRRYKGSIFYKDMTGTFTLDANTL-KRASRHVF 197
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Qy 106 KWTGK--GMLLAGAVTMAVTRLTEIIPFTFANSYN-RKLKNSLNIAFGGFPSPFD--- 159
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 198 RKSSGLVIGMELGGSTWFSASNLNLT-----PFNQVKSRITFQLQGRFGVWNDEYDIDRYG 253
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 160 --INMQASA-LGPEL-SFKKSY 178
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 254 DEIVLGGSSVELGVKVPKFNKY 276
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :

RESULT 11
US-10-156-761-11883
; Sequence 11883, Application US/10156761
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
```

```
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-282
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11883
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11883

Query Match 9.0%; Score 81; DB 6; Length 379;
Best Local Similarity 24.2%; Pred. No. 3.5;
Matches 37; Conservative 26; Mismatches 50; Indels 40; Gaps

Qy 27 AFNLGAGEKLL----AYETSKKDPVPPFLLNLFGLGIGSFAQGDILGGLILGFDVAVGI 82
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 240 ALALGLGEGVLGTSAY-TSKREQF-----GFPVATFQAVAV 275
|| : | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Qy 83 GLILTGAYLDIKALDKNAKPAKFKWTGKGMLAGAVTMAVTRLTEIIPFTFANSYNRK 142
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 276 Q--AADRYIDLKAMEATLWQAARISTGAGGALPAAGDVAVAKI-----WASEGVRR 325
: | | : | | : | | : | | : | | : | | : | | : | | : | | : | | :
Qy 143 LKNSLNIAFGGF--EPSFDINMQASALGFELS 173
: | | | : | | | : | | | : | | | : | | | : | | | : | | | :
Db 326 VVQTAQLHLGGFGADTDYPLHRYHAWAKHLELS 358
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

RESULT 12
US-10-282-122A-51378
; Sequence 51378, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
```



```

; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66257
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66257

```

```

Query Match      8.7%; Score 78.5; DB 6; Length 436;
Best Local Similarity 37.1%; Pred. No. 7.4;
Matches 26; Conservative 7; Mismatches 24; Indels 13; Gaps 3;

QY  22 DDSKSAFNIGAGKLLAYETSKDPVFPFLNLFGLGIGSGFAOGDILGGFLILGFDVAV- 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  114 DTSKIAAYAAAG-----KSTVDFLMNVIPGTVVGAFANGDIL---QVLFFSVLF 161

QY  81 GIGLILTGAY 90
    ||||| |||||
Db  162 GYALHRLGSY 171

```

Search completed: March 14, 2003, 14:36:26
Job time : 25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2003, 20:13:37 ; Search time 29.3911 Seconds
(without alignments)
7815.311 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749

Sequence: 1 tagaattttcaacaataaaa.....atttatctagaataatgggtg 749

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	51.4	6.9	20674	4	US-09-641-638-651
2	49.6	6.6	665	2	US-08-883-795A-36
3	48.6	6.5	837	4	US-08-998-416-288
4	47.2	6.3	665	2	US-08-883-795A-36
5	46.8	6.2	2251	4	US-08-991-677-11
6	45.2	6.0	19124	2	US-08-487-826B-13
7	44.4	5.9	8920	2	US-08-446-855A-1
8	44.4	5.9	8920	4	US-09-150-741-1
9	44.2	5.9	5852	1	US-07-867-106-2
10	42.8	5.7	615	4	US-08-998-416-186
11	42.8	5.7	636	4	US-08-998-416-1137
12	42.8	5.7	837	4	US-08-998-416-288
13	42.8	5.7	1850	3	US-08-617-860B-32
14	42.8	5.7	3440	1	US-08-471-791-27
15	42.8	5.7	3440	5	PCT-US91-01746-27
16	42.8	5.7	4098	2	US-08-605-106-4
17	42.6	5.7	1850	3	US-08-617-860B-32
18	42.6	5.7	4098	2	US-08-605-106-4
19	42	5.6	731	1	US-08-451-405A-2
20	42	5.6	834	4	US-08-998-416-305
21	41.4	5.5	2960	3	US-08-913-842-3
22	41.2	5.5	2251	4	US-08-991-677-11
23	41	5.5	2663	1	US-08-136-743B-3
24	40.6	5.4	767	4	US-08-998-416-472
25	40.4	5.4	782	4	US-08-998-416-224
26	40.4	5.4	1454	2	US-08-713-000-7
27	40.4	5.4	1454	2	US-08-975-316-7

c 28	40.4	5.4	1454	4	US-09-211-710-7	Sequence 7, Appli
c 29	40.4	5.4	1454	4	US-09-615-192A-7	Sequence 7, Appli
c 30	40.4	5.4	1460	4	US-09-615-192A-95	Sequence 95, Appl
c 31	40.4	5.4	1474	2	US-08-975-316-71	Sequence 71, Appl
c 32	40.4	5.4	1474	4	US-09-615-192A-71	Sequence 71, Appl
c 33	40.4	5.4	1474	4	US-09-615-192A-402	Sequence 402, App
c 34	40.4	5.4	1474	4	US-09-325-932A-202	Sequence 202, App
c 35	40.4	5.4	5303	4	US-08-971-395-4	Sequence 4, Appli
c 36	40.4	5.4	6124	4	US-08-213-419B-3	Sequence 3, Appli
c 37	40.4	5.4	6124	4	US-08-213-419B-3	Sequence 3, Appli
c 38	40.4	5.4	6243	2	US-09-056-075-1	Sequence 1, Appli
c 39	40	5.3	1422	1	US-08-319-704-5	Sequence 5, Appli
c 40	40	5.3	6243	2	US-09-056-075-1	Sequence 1, Appli
c 41	39.8	5.3	600	4	US-09-134-001C-1770	Sequence 1770, Ap
c 42	39.8	5.3	636	4	US-08-998-416-1137	Sequence 1137, Ap
c 43	39.8	5.3	2672	1	US-08-703-947-1	Sequence 1, Appli
c 44	39.8	5.3	3850	3	US-08-329-799-34	Sequence 34, Appl
c 45	39.6	5.3	662	4	US-08-998-416-185	Sequence 185, App

ALIGNMENTS

RESULT 1

US-09-641-638-651

; Sequence 651, Application US/09641638

; Patent No. 6432648

; GENERAL INFORMATION:

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Bougueleret, Lydie

; APPLICANT: Chumakov, Ilya

; APPLICANT: Cohen, Annie

; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING

; FILE REFERENCE: GENSET.051CPI

; CURRENT APPLICATION NUMBER: US/09/641.638

; PRIOR FILING DATE: 2000-08-16

; PRIOR APPLICATION NUMBER: US 09/502,330

; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: US 60/133,200

; PRIOR FILING DATE: 1999-05-07

; PRIOR APPLICATION NUMBER: US 09/275,267

; PRIOR FILING DATE: 1999-03-23

; PRIOR APPLICATION NUMBER: US 60/119,917

; PRIOR FILING DATE: 1999-02-12

; NUMBER OF SEQ ID NOS: 1304

; SOFTWARE: Patent.pm

; SEQ ID NO 651

; LENGTH: 20674

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 1123..3123

; OTHER INFORMATION: 5'regulatory region

; NAME/KEY: exon

; LOCATION: 3124..3297

; OTHER INFORMATION: exon 1

; NAME/KEY: exon

; LOCATION: 3871..4072

; OTHER INFORMATION: exon 2

; NAME/KEY: exon

; LOCATION: 5552..5633

; OTHER INFORMATION: exon 3

; NAME/KEY: exon

; LOCATION: 5758..5880

; OTHER INFORMATION: exon 4

; NAME/KEY: exon

; LOCATION: 5996..6099

; OTHER INFORMATION: exon 5

; NAME/KEY: exon

; LOCATION: 6349..6509

; OTHER INFORMATION: exon 6

```
NAME/KEY: exon
LOCATION: 7379..7522
OTHER INFORMATION: exon 7
NAME/KEY: exon
LOCATION: 8645..8854
OTHER INFORMATION: exon 8
NAME/KEY: exon
LOCATION: 12254..12340
OTHER INFORMATION: exon 9
NAME/KEY: exon
LOCATION: 12854..13023
OTHER INFORMATION: exon 10
NAME/KEY: exon
LOCATION: 13308..13429
OTHER INFORMATION: exon 11
NAME/KEY: exon
LOCATION: 16567..16667
OTHER INFORMATION: exon 12
NAME/KEY: exon
LOCATION: 16775..16945
OTHER INFORMATION: exon 13
NAME/KEY: exon
LOCATION: 17063..17554
OTHER INFORMATION: exon 14
NAME/KEY: misc_feature
LOCATION: 17555..20674
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 1128
OTHER INFORMATION: 10-508-191 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1182
OTHER INFORMATION: 10-508-245 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1559
OTHER INFORMATION: 10-509-284 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1570
OTHER INFORMATION: 10-509-295 : deletion of C
NAME/KEY: allele
LOCATION: 1827
OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTTTT
NAME/KEY: allele
LOCATION: 2048
OTHER INFORMATION: 10-511-62 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2323
OTHER INFORMATION: 10-511-337 : insertion of T
NAME/KEY: allele
LOCATION: 2341
OTHER INFORMATION: 10-512-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 2623
OTHER INFORMATION: 10-512-318 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2832
OTHER INFORMATION: 10-513-250 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2844
OTHER INFORMATION: 10-513-262 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2934
OTHER INFORMATION: 10-513-352 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2947
OTHER INFORMATION: 10-513-365 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 3802
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6375
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6429
OTHER INFORMATION: 10-347-165 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6467
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6484
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6534
OTHER INFORMATION: 10-347-271 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 6611
OTHER INFORMATION: 10-347-348 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 7668
OTHER INFORMATION: 10-348-391 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8608
OTHER INFORMATION: 10-349-47 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 8658
OTHER INFORMATION: 10-349-97 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8703
OTHER INFORMATION: 10-349-142 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 8777
OTHER INFORMATION: 10-349-216 : deletion of CTG
NAME/KEY: allele
LOCATION: 8785
OTHER INFORMATION: 10-349-224 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 8926
OTHER INFORMATION: 10-349-368 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12171
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12429
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13341
OTHER INFORMATION: 10-507-170 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 13492
```


NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG124LRP
US-08-998-416-288

Query Match 6.5%; Score 48.6; DB 4; Length 837;
Best Local Similarity 46.6%; Pred. No. 0.017; Mismatches 179; Indels 0; Gaps 0;
Matches 156; Conservative 0;
QY 6 TTTTCAACAAATAAGATATGTTAAAGAAATGCTAAATTTTATGGTTAAATCA 65
DB 611 TTTTAAATTTTATTAATTAATTAATATGATAAAATATTAATTAATTAATTA 552
QY 66 AGAGCTCTATGGTGAAGCAATTCGAGTAACAATTTGAAAGAGTTAAATTAATAG 125
DB 551 ATTAAATAAATAAATAAATAAATAAATAAATTTATTAAATAAATAAATAAATAA 492
QY 126 TTCCAAAGCCCTTTTAAATTTTCAATTAATATGCTACCATAATACCAGTTTAAATAGG 185
DB 491 TAAGAAATTAAGTTAAATTAATTTTAAATAATATCTTATAAAGAGTTAAATATAT 432
QY 186 GTTTTATGAATAAGTTTTTAAATTTTATTTTGGTAACTTTTGTGCTTTTCTAGTTT 245
DB 431 AAATCAACATAATATTTATAAAATAGATATATATAAATAAATAAATAATATTT 372
QY 246 GCTCAGATGATCTAAAGCAGCTTTTAACTGGGAGCGGAGAGAAAATTTTGGTTAT 305
DB 371 AAATAAATTAATCTTTATAAATAAATAAATAATATTTTATAAACAATAAATAAATA 312
QY 306 GAAACTAATAAGAGATCTCTGTGACCATTTT 340
DB 311 TATATTAATTTGATAATCTATTAATTAATTTAT 277

RESULT 4
US-08-795A-36/C
Sequence 36, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcuve, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883.795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: Rh 32
US-08-883-795A-36
Query Match 6.3%; Score 47.2; DB 2; Length 665;
Best Local Similarity 45.8%; Pred. No. 0.032;
Matches 163; Conservative 0; Mismatches 193; Indels 0; Gaps 0
QY 1 TAGAATTTTCACAAATAAGATATGTTAAAGAAATTCGAAATGCTAAATTTATGTTA 60
DB 401 TAAACATTTTAAATATATAAATATGTAATTAATAACATTTTAAATATAAAATATGTAATTA 342
QY 61 AATCAAGAAGCTCTATTTGGTAAAGCAATTCGAGTAGCACAATTTGAAAAAGTTAAATTTA 120
DB 341 TAAACATTTTAAATATATAAATATGTAATTAATAACATTTTAAATATAAAATATTTAATTA 282
QY 121 ATAGTTCCAAAGCCCTTTTAAATTTTCAATTAATATGTCACATTAATACCAGTTTAAATA 180
DB 281 TAAACATTTTAAATATAAATAATTTTAAATATAAATAATTTTAAATATAAATAATTTAATTA 222
QY 181 AAGGGTTTTTATGATAGTATTTTAAATTTTATTTTGGTAACTTTTGTGCTTTTCTA 240
DB 221 TAAATATTTTAAATATAAATAATTTTAAATATAAATAATTTTAAATATAAATAATTTAATTA 162
QY 241 GTTTTGCTCAAGATGATTTCTAAAGCACATTTTAAATCTGGAGCGGAGAGAAAATTTTGG 300
DB 161 TAAATATTTTAAATATAAATAATTTTAAATATAAATAATTTTAAATATAAATAATTTAATTA 102
QY 301 TTTATGAACTAATAAGAAAGATCTCTGTGACCATTTTATTTGAAACCTTTTAAATTA 356
DB 101 TAAATATTTTAAATATAAATAATTTTAAATATAAATAATTTTAAATATAAATAATTTAATTA 46

RESULT 5
US-08-991-677-11
Sequence 11, Application US/08991677A
Patent No. 6252135
GENERAL INFORMATION:
APPLICANT: Chiang, Vincent L
APPLICANT: Carriway, Daniel T
APPLICANT: Smeltzer, Richard H
TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
FILE REFERENCE: 50617
CURRENT APPLICATION NUMBER: US/08/991,677A
CURRENT FILING DATE: 1997-12-16
EARLIER APPLICATION NUMBER: US 60/033,381
EARLIER FILING DATE: 1996-12-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 2251
TYPE: DNA
ORGANISM: Pinus taeda
US-08-991-677-11

Query Match 6.2%; Score 46.8; DB 4; Length 2251;
Best Local Similarity 52.6%; Pred. No. 0.052; Mismatches 92; Indels 0; Gaps 0;
Matches 102; Conservative 0;

Db 2261 A 2261

RESULT 10

```

RESULT 10
US-08-998-416-186
: Sequence 186, Application US/08998416
: Patent No. 6239264
: GENERAL INFORMATION:
: APPLICANT: Philippsen, Peter
: APPLICANT: Pohlmann, Rainer
: APPLICANT: Steiner, Sabine
: APPLICANT: Mohr, Christine
: APPLICANT: Wendland, Jorgen
: APPLICANT: Knechtle, Philipp
: APPLICANT: Rebschung, Corinne
: TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
: TITLE OF INVENTION: AND USES THEREOF
: NUMBER OF SEQUENCES: 1152
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NO. 6239264artis Corporation
: STREET: 3054 Cornwallis Road
: CITY: Research Triangle Park
: STATE: No. 6239264th Carolina
: COUNTRY: USA
: ZIP: 27709
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/998,416
: FILING DATE: 24-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: CH 0016/97
: FILING DATE: 31-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8587
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 186:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 615 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: PAG1074RP
US-08-998-416-186

```

Query Match 5.7%; Score 42.8; DB 4; Length 615;
Best Local Similarity 46.9%; Pred. NO. 0.3;
Matches 167; Conservative 0; Mismatches 187; Indels 2; Gaps 1;

QY	1	TAGAATTTTCACCAATAAAGATATTGTTAAAAAGAAATCGAAATTCGTAAATTTTATGCTTA	60
Db	39	TAAACTTTTTTTATTATAAATTTTAACTATTAATTTTAACTATTAATATCATATTATTT	98
QY	61	ANTCAGAAGCTCTATTGGTGAACGGAATTCGAGTACAATTTGAAAAAGTTA--AATT	118
Db	99	AATAAAATTAATTTATTCATTATTAATACTATTATATAATTTATTTATTAATTTACTTAAT	158
QY	119	TAAATAGTCCAAAGCCCTTTTTTAATTTTCATTAATATGCTACCATAATACCAGTTTAA	178
Db	159	TCATCATATTATATATTTATATAAATTTATTAATTAATTAATTAATTAATTAATTAATTAAT	218
QY	179	TAAAGGGGTTTTTATGAATAAGTTTTTAATTTTTTATTTTGGTAAATCTTTTGCTTTTTC	238

```

Db      219  CTATGTTCAAAATTTTAAATAGTATTAAAAATATTATTAGATATTATTATTTCTCTTTAAT 278
Qy      239  TAGTTTTGCTCAAGAGTATTCTAAAAGCACCTTTTAAATCTGGGAGCGGAGAAAAATTTTT 298
Db      279  AAATTATTAAATAGATATATCAATTAATTAATATATTATTATTAAATGTTTATTAATAATAA 338
Qy      299  GGTTTATGAACATAAATAAGAAAGATTCTCTGTGTACCAATTTTATGTAAACCTTTTTTT 354
Db      339  TATATTTTATTATTATAAAGATTAAATTTATTTAAAAATATTGTAATTAATTTATTTT 394

RESULT 11
US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebeschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709

```

RESULT 11

```

US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgon
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAGI692RP
; US-08-998-416-1137

```

Query Match	5.7%	Score 42.8;	DB 4;	Length 636;
Best Local Similarity	46.9%	Pred. No. 0.3;		
Matches 167;	Conservative	0;	Mismatches 187;	Indels 2;
Gaps 1;				

Qy	1	TAGAAATTTTCAACAAATAAAGATATCTGTTAAAGAAATTTGAAATTTGCTAAATTTTATGGTTA	60
Db	39	TAAACTTTTATTATTAATAATTAAAGTATTAAATATTTTAACTATTATTATCATTAATTT	98
Ov	61	AATCAAGAAGCTCTATTGGTAAAGCGAATTTTCGACTAACAAATTTGAAAAGAAAGCTTA--AATT	111


```

RESULT 15
PCT-US91-01746-27/c
; Sequence 27, Application PC/TUS9101746
; GENERAL INFORMATION:
; APPLICANT: Thompson, Gregory A
; APPLICANT: Knauf, Vic C
; TITLE OF INVENTION: Plant Desaturases-Compositions and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: California
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.7
; SOFTWARE: Microsoft Word 4.0

```

Search completed: March 15, 2003, 07:44:33
Job time : 58.3911 secs

	Qy	MNKFLLVLLAFVCVFSSFAQAADSKSAFNLGAGEKLLAYETSKKDPIVPFLNFLFGFI	60
	Dd	MNKFLLVLLAFVCVFSSFAQAADSKSAFNLGAGEKLLAYETSKKDPIVPFLNFLFGFI	60

QY 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKSY 178
|||||
Db 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKSY 178
|||||

RESULT 2

US-09-156-447-21
; Sequence 21, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-156-447-21

Query Match 100.0%; Score 904; DB 15; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.8e-92;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
Db 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
QY 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKSY 178
|||||
Db 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKSY 178
|||||

RESULT 3

US-09-508-487-21
; Sequence 21, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-508-487-21

Query Match 100.0%; Score 904; DB 19; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.8e-92;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
Db 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
QY 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKSY 178
|||||
Db 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKSY 178
|||||

RESULT 4

US-09-153-447-23
; Sequence 23, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 60/059,036
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 1041/97
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-153-447-23

Query Match 87.0%; Score 786.5; DB 15; Length 177;
Best Local Similarity 86.5%; Pred. No. 2.6e-79;
Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVLPVLLNLFGLFGI 60
|||||
Db 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKKDPVLPVLLNLFGLFGI 59
|||||
QY 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
|||||
Db 61 GSPAQGDILGGFLILGFDVAVGIGLILTGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 119
|||||
QY 121 MAVTRLTEIIIPFTFANSYNRKLNLSNIAFGGFEPSFDINMGQASALGFELSFKKSY 178
|||||
Db 120 MAVTRLTEIVLPFTFANNYNRKLNLSNIALGGFEPSFDINMGQASALGFELSFKKSY 177
|||||

RESULT 5

US-09-156-447-23
; Sequence 23, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-156-447-23

```
Query Match      87.0%; Score 786.5; DB 15; Length 177;
Best Local Similarity 86.5%; Pred. No. 2.6e-79;
Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLGFI 60
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 1 MNKFLIFILVIFCAFSSFAQ-DDSKSTFNLGAGEKFLVYETNKKKDSLVPELLNLFGLGFI 59
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 61 GSFAQGDILGSLILGFDVAGIGLITGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 60 GSFAQGDILGSLILGFDVAGIGLITGAYLDIKDFDNNAKKADFKWTWCKGMMLAGVVT 119
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 121 MAVTRLTETIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSPKKS 178
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 120 MAVTRLTETIIPFTFANNYNRKLNLSNIALGGPEPSFDINMGQASALGFELSPKKS 177
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||

RESULT 6
US-09-508-487-23
; Sequence 23, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130
; CURRENT APPLICATION NUMBER: US/09/508,487
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: PCT/IB98/01424
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Borrelia garinii
US-09-508-487-23

Query Match      87.0%; Score 786.5; DB 19; Length 177;
Best Local Similarity 86.5%; Pred. No. 2.6e-79;
Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLGFI 60
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 1 MNKFLIFILVIFCAFSSFAQ-DDSKSTFNLGAGEKFLVYETNKKKDSLVPELLNLFGLGFI 59
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 61 GSFAQGDILGSLILGFDVAGIGLITGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAVT 120
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 60 GSFAQGDILGSLILGFDVAGIGLITGAYLDIKDFDNNAKKADFKWTWCKGMMLAGVVT 119
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 121 MAVTRLTETIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSPKKS 178
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 120 MAVTRLTETIIPFTFANNYNRKLNLSNIALGGPEPSFDINMGQASALGFELSPKKS 177
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||

RESULT 7
US-09-153-447-19
; Sequence 19, Application US/09153447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/153,447
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 60/059,036
; EARLIER FILING DATE: 1997-09-10
; EARLIER APPLICATION NUMBER: 1041/97
; EARLIER FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
```

```
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-153-447-19

Query Match      85.1%; Score 769.5; DB 15; Length 179;
Best Local Similarity 84.9%; Pred. No. 2.1e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLGFI 59
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 1 MNKLLIEVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPELLNLFGLGFI 60
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 60 IGSAQGDILGSLILGFDVAGIGLITGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAV 119
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 61 IGSAQGDILGSLILGFDVAGIGLITGAYLDIKALDITKKAFOFTWCKGMMLAGAV 120
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 120 TMAYTRLTETIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSPKKS 178
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 121 TMAYTRLTETIIPFTFANSYNRKLNLSNIALGGPEPSFDVAMGQSSALGFELSPKKS 179
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||

RESULT 8
US-09-156-447-19
; Sequence 19, Application US/09156447
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3130.1
; CURRENT APPLICATION NUMBER: US/09/156,447
; CURRENT FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Borrelia burgdorferi
US-09-156-447-19

Query Match      85.1%; Score 769.5; DB 15; Length 179;
Best Local Similarity 84.9%; Pred. No. 2.1e-77;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPIVPELLNLFGLGFI 59
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 1 MNKLLIEVLATFCVFSFAQANDSKNGAFGMSAGEKLLVYETSKQDPIVPELLNLFGLGFI 60
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 60 IGSAQGDILGSLILGFDVAGIGLITGAYLDIKALDKNAPKAAFKWTWCKGMMLAGAV 119
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 61 IGSAQGDILGSLILGFDVAGIGLITGAYLDIKALDITKKAFOFTWCKGMMLAGAV 120
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
QY 120 TMAYTRLTETIIPFTFANSYNRKLNLSNIAFGGPEPSFDINMGQASALGFELSPKKS 178
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||
Db 121 TMAYTRLTETIIPFTFANSYNRKLNLSNIALGGPEPSFDVAMGQSSALGFELSPKKS 179
      ||||| : : || ||||| ||||| ||||| : ||||| : ||||| |||||

RESULT 9
US-09-508-487-19
; Sequence 19, Application US/09508487
; GENERAL INFORMATION:
; APPLICANT: BERGSTROM, Sven
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
; FILE REFERENCE: 454312-3150
; CURRENT APPLICATION NUMBER: US/09/508,487
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/059,036
; PRIOR FILING DATE: 1997-09-10
; PRIOR APPLICATION NUMBER: 1041/97
; PRIOR FILING DATE: 1997-09-10
```

```

RESULT 13
US-09-791-537-91922
; Sequence 91922, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL
; TITLE OF INVENTION: METHODS OF USE TH
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
SEQ ID NO 91922

```

Qy 41 TSKKDPVPPFLNLFGLFGIGSFAQ 65
Db 1 TSKQDPVPPFLNLFGLFGIGSFAQ 25

Search completed: March 14, 2003, 14:35:07
Job time : 146.333 secs

LENGTH: 95
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-791-537-91922

Query Match 17.7%; Score 160; DB 21; Length 95;
Best Local Similarity 44.4%; Pred. No. 2.1e-09;
Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;

Qy 1 MNKFLIVLLAFVFPSSFAQADDSKAFNLGAGEKLLAYETSKKDPVPPFLNLFGLFGI 60
Db 1 MKKIPTLILFGLTIEIFA-TKDTQNIKEGI-ESFNKYDKREKNPIGPFLLNLFLEFGI 58

Qy 61 GSFAQGDILGGFLILGFDVAG 81
Db 59 GSFVQGDYIGGSGVLGFNLG 79

RESULT 14
US-09-153-447-1
Sequence 1, Application US/09153447
GENERAL INFORMATION:
APPLICANT: BERGSTROM, Sven
TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
FILE REFERENCE: 454312-3130.1
CURRENT APPLICATION NUMBER: US/09/153.447
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: 60/059,036
EARLIER FILING DATE: 1997-09-10
EARLIER APPLICATION NUMBER: 1041/97
EARLIER FILING DATE: 1997-09-10
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 25
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-153-447-1

Query Match 13.7%; Score 124; DB 15; Length 25;
Best Local Similarity 96.0%; Pred. No. 4.3e-06;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 41 TSKKDPVPPFLNLFGLFGIGSFAQ 65
Db 1 TSKQDPVPPFLNLFGLFGIGSFAQ 25

RESULT 15
US-09-156-447-1
Sequence 1, Application US/09156447
GENERAL INFORMATION:
APPLICANT: BERGSTROM, Sven
TITLE OF INVENTION: P13 ANTIGENS FROM BORRELIA
FILE REFERENCE: 454312-3130.1
CURRENT APPLICATION NUMBER: US/09/156.447
CURRENT FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/059,036
PRIOR FILING DATE: 1997-09-10
PRIOR APPLICATION NUMBER: 1041/97
PRIOR FILING DATE: 1997-09-10
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 25
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-156-447-1

Query Match 13.7%; Score 124; DB 15; Length 25;
Best Local Similarity 96.0%; Pred. No. 4.3e-06;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)


```
Db 179 LPLFIQIFLQGM-----NGLLEIMSAFGKVLILA--VGTLLFLAIQIIFIAGAVSKKN 232
QY 101 PKAAFKWTWKGMMLAGAVMTVRLTEIIPFTFANSYNRKLNLSNIAFGGPEP----- 156
Db 233 PWKLFK-----NMLPAYFALGFSSSAAATPVT-----YQOTLKNDVDVNVAGFVPLCA 282
QY 157 -----SFDINMGQASALGFEL 172
Db 283 TIHLAGSMNKIGLFTFAVFMIDMEVGVGLSIGFL 318

RESULT 2
US-09-815-242-13341
; Sequence 13341, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13341
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13341

Query Match 9.2%; Score 83.5; DB 10; Length 650;
Best Local Similarity 25.3%; Pred. No. 0.97;
Matches 39; Conservative 22; Mismatches 54; Indels 39; Gaps 7;

QY 32 AGEKLLAYERSKKDP-----IVPFLNLFLG-----GIGSFAOGDILGGF 72
Db 420 AGALVLAIKYVVRPSLEGAKSILLPLLTGTILTGFMVLAVNIPMAINTANDEF LG- 478
QY 73 LILGFDVAGTGLITGAYLIDKALDNAPKAFKWTWKGMMILA-----GAVTMA----- 122
Db 479 -LGGSAVLIGVILGG-----MMAVDMPGVNKAAYVFGTGLTAATVSSGGSVAMAAVMAG 533
QY 123 --VTRLTEIIPFTFANSYNRKLNLS--LNIAFG 152
Db 534 GMVPPPLAIFVATLLFKDKFTKEERNISGLTNIIMG 567

RESULT 3
US-09-910-186A-22
; Sequence 22, Application US/09910186A
; Publication No. US2003009025A1
```

```
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Research & Material Command
; TITLE OF INVENTION: RECOMBINANT VACCINE AGAINST BOTULINUM
; FILE REFERENCE: NEUROTOXIN
; FILE REFERENCE: A33626-A 067252.0107
; CURRENT APPLICATION NUMBER: US/09/910,186A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: PCT/US00/12890
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 09/611,419
; PRIOR FILING DATE: 2000-07-06
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,865
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,866
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,867
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,868
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,869
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,873
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 08/123,975
; PRIOR FILING DATE: 1993-09-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-09-910-186A-22

Query Match 9.0%; Score 81.5; DB 9; Length 413;
Best Local Similarity 23.8%; Pred. No. 0.89;
Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;

QY 47 IVPFLNLFLGFGIGSFAOGDILGGFLILGFDV-----GIGLITGAYLD----- 92
Db 180 IVPY-IGLANVVG-NETAKGNFENAFIAGASILLEFIPELLIPVVGAFLLSEYIDNKK 237
QY 93 -IKALDNAPKAFKWTWKGMMLAGAVTMAVTRL-----TIIIPFT 134
Db 238 IIKTDNALTKRNEKWSMDYGLIVAOWLSTVNTQFYIKGMYKALNYQAQALEEIKYR 297
QY 135 FANSYNRKLKNSLNIAFGGPEPSFDINNGOA 165
Db 298 Y-NIYSEKEKSNINIDFNDSKLNNEGINOQ 327

RESULT 4
US-10-011-588-25
; Sequence 25, Application US/10011588
; Patent No. US20020168727A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Leonard
; APPLICANT: Jensen, Melody
; TITLE OF INVENTION: RECOMBINANT LIGHT CHAINS OF BOTULINUM
; TITLE OF INVENTION: NEUROTOXINS AND LIGHT CHAIN FUSION PROTEINS FOR USE IN
; FILE REFERENCE: RESEARCH AND CLINICAL THERAPY
; FILE REFERENCE: A34796 067252.0113
; CURRENT APPLICATION NUMBER: US/10/011,588
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/910,186
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/611,419
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/246,744
; PRIOR FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: 60/311,966
; PRIOR FILING DATE: 2001-08-09
```

NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant protein encoded by SEQ ID NO:24
US-10-011-588-25

Query Match 9.0%; Score 81.5; DB 9; Length 852;
Best Local Similarity 23.8%; Pred. No. 2.2;
Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;
Qy 47 IVPFLLNLFGLFGSGFAQGDILGGLILGFDV-----GIGLILGAYLD----- 92
Db 619 IVPY-IGLALNGV-NETAKGFENAFETAGASILLEFTELLIPVVGAFLESYIDNKNK 676
Qy 93 -IKALDKNAPKAAFKWTGKGMLAGAVTMAVTRL-----TEIIIPFT 134
Db 677 IIKTIDNALTNRNEKWSMDYGLIWAQWLSVTNTQFYTIKEGMYKALNYQAQALEIIKYR 736
Qy 135 FANSYNRKLKNSLNIAFGFSPSPDINMQQA 165
Db 737 Y-NIYSEKESKNINIDFNDINSKLNQGINQA 766

RESULT 5
US-09-815-242-10070
; Sequence 10070, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10070
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(663)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-14080

Query Match 8.7%; Score 78.5; DB 10; Length 663;
Best Local Similarity 22.6%; Pred. No. 3.3;
Matches 38; Conservative 19; Mismatches 72; Indels 39; Gaps 6;
Qy 28 FNLGAGEKLLAYETSKDPPIVFFLLNLF-----LGFGISFAQGDILG----- 70
Db 118 FVIGLMLNVPLQIGARDVAFPPFNLSFWFVVGVLVNLISLGVGEFAQTGLAYPPPLS 177
Qy 71 -----GFLILGFDVAVGIGLILTCAYLDIKALDKNAP-KAAFK----WTW----GKG 112
Db 178 GIEYSPSGVDYWIWALQSLGIGTTLTGINFVFTILKMRAPGTMFKMPVFTWASLCANV 237
Qy 113 MMLAG-----AVTMAVTRLTEIIIPFTFANSYNRKLKNSLNIAFGGFEP 156
Db 238 LIIASFPILTVTVALLTDRLYLGHFFFTNDMGNNMMYINLIWAGHP 285

RESULT 7
US-09-815-242-10070
; Sequence 10070, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10070
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10070

Query Match 8.7%; Score 78.5; DB 10; Length 663;
Best Local Similarity 22.6%; Pred. No. 3.3;
Matches 38; Conservative 19; Mismatches 72; Indels 39; Gaps 6;
Qy 28 FNLGAGEKLLAYETSKDPPIVFFLLNLF-----GIGSFAQGDILG----- 70
Db 118 FVIGLMLNVPLQIGARDVAFPPFNLSFWFVVGVLVNLISLGVGEFAQTGLAYPPPLS 177
Qy 71 -----GFLILGFDVAVGIGLILTCAYLDIKALDKNAP-KAAFK----WTW----GKG 112
Db 178 GIEYSPSGVDYWIWALQSLGIGTTLTGINFVFTILKMRAPGTMFKMPVFTWASLCANV 237
Qy 113 MMLAG-----AVTMAVTRLTEIIIPFTFANSYNRKLKNSLNIAFGGFEP 156
Db 238 LIIASFPILTVTVALLTDRLYLGHFFFTNDMGNNMMYINLIWAGHP 285

•

```
Best Local Similarity 23.6%; Pred. No. 1.6;
Matches 33; Conservative 20; Mismatches 38; Indels 49; Gaps 6;

Qy 61 GSFAQGDILGFLILGFDVAGIGLILTCAYLDIKALDNKAPKAFKWTWG-----K 111
Db 86 GHFVOGHI-----DAIGV-----IEKIHNANQVDFISASEETLLLCVEQ 126
Qy 112 GMLAGAVTMAVTRLTE-----IIIPFTANSYNR--KLKNSLNTAFG----- 152
Db 127 GSIADVGSVLTLSKVEKGFWLTIPTLENTLFKAYLKRVRNLETDMLVRSVASILKK 186
Qy 153 --GFEPFDMNQASALGF 170
Db 187 TKGEKFNFSWEADALTLY 206

RESULT 10
US-08-910-386A-20
; Sequence 20, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-910-386A-20

Query Match 8.1%; Score 73.5; DB 8; Length 1161;
Best Local Similarity 28.4%; Pred. No. 22;
Matches 31; Conservative 17; Mismatches 42; Indels 19; Gaps 5;

Qy 76 GFDAVGILTLTGAVLDIKALDNKAPKAFKWT--WCKGMLAGAVTMAVTRLTEIIPF 133
Db 830 GFNSIETASITSAAIYVSLV---ALIVLFYTRKNWRSRVAGSTRKEVTFTEVPVL 886
Qy 134 TFAN-----SYNRKLKNSLNTAFGFGFEPFDMNQASALGFELSFKK 176
Db 887 TFENVVRATGSFNA-----SNCIGSGGFGATYKAEI----APGFLVAVKR 927
```

```
RESULT 11
US-09-741-669-422
; Sequence 422, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; TITLE OF INVENTION: proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741.669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 422
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-741-669-422

Query Match 8.0%; Score 72.5; DB 10; Length 451;
Best Local Similarity 26.2%; Pred. No. 8.7;
Matches 34; Conservative 14; Mismatches 45; Indels 37; Gaps 7;

Qy 46 PIVFPLMLFLGFGIGSPAQGDILGGLILGFDVAGIGLILTGAYLD-----IKAL----- 96
Db 19 PIVIIIFSKILG-----MKAGDCFKAGLHIGIFGVIGLVI-GLMLDSIGSPAAMAEKF 72
Qy 97 DKNAPKAAFKW-----TWGKGMLAGA-----VTMAVTRLTEIIPFTFANSYNRKL 143
Db 73 DLNLHVVDVGPWGSSPMTWASQIALVPIALVNVAMLLTRMTKV-----NVDI 123
Qy 144 KNSLNIAPFG 153
Db 124 WNIWMTFTG 133

RESULT 12
US-09-925-637-16
; Sequence 16, Application US/09925637
; Patent No. US20020103338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus polynucleotides and polypepti
; FILE REFERENCE: PB560
; CURRENT APPLICATION NUMBER: US/09/925.637
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-637-16

Query Match 8.0%; Score 72; DB 10; Length 452;
Best Local Similarity 21.3%; Pred. No. 9.9;
Matches 37; Conservative 28; Mismatches 71; Indels 38; Gaps 6;

Qy 27 AFNLGA-GEKLLAYETSKDFIVPFLNLFLGFGI---GSFAQGDILGFLILGFDVAGI 82
```

Db 24 AFLGRYGGVLAHNGKHP-----RVLVGRDTRVSGEMLESALIAGLISICAEVMRL 77
QY 83 GLILT-----GAYLDIKALDKNAPKA--AFKWTWCKMMLAGAVTMVTRLTEII 130
Db 78 GIISTPGVAYLTRDMGAELGVMISASHNPVADNGIKFFGSDGPKLSDEQNEIEALLDQE 137
QY 131 IP-----FTFANSYNRKLNLSNIAFGGFEPSFDINMGQASAL 168
Db 138 NPELPRPVGNDIVHYSIDYFEGAQKYLKSTVDVNFEGLIKIALDGANGSTSSL 191

RESULT 13
US-09-881-752A-16
; Sequence 16, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: KleanChous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Omen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-16

Query Match 7.9%; Score 71.5; DB 10; Length 220;
Best Local Similarity 26.1%; Pred. No. 4.6;
Matches 47; Conservative 20; Mismatches 62; Indels 51; Gaps 10;
QY 1 MNKFLIIVLLAFVCFVSFAQD---DSKS-----AFNLGAGEKLLAYETSKDPIV 48
Db 1 MNKTTVKILMGMAALLSQAEEADELDEKSKPKFADNRTFYLGVGQLSAINTSFSTESV 60
QY 49 PFLNLFL---GRGIGSFAQGDILGFLGFDV-----GIGLILT-----GAYLDIK 94
Db 61 D--KSYFMTGNGRGV-----VLGKFKVAKTQAVEHVGFRYGLFYDQTFSSHKSIISY 111
QY 95 ALD-----KNAPKAAFKWTWCKMMLAGAVTMVTRLTEIIPFTFANSYNRKLNLSL 147
Db 112 GLEFGSLWDAFNPKPKFLGLEFLG---IAGATYMPGGAMHGII-----AQNLGKENS 162

RESULT 14
US-09-738-626-4426
; Sequence 4426, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 4426
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4426

Query Match 7.9%; Score 71.5; DB 9; Length 292;
Best Local Similarity 26.5%; Pred. No. 6.5;
Matches 43; Conservative 18; Mismatches 56; Indels 45; Gaps 9;
QY 18 FAQADDSKSAF-----NLGAGEKLLAYETSKDPIVPELNLFLGFGISFAQGDIL-- 69
Db 18 FVSASDPQSIIRSEPRADRCYGRKLLA-QLNPTWPTP-----IQGFALNRSVPA 66
QY 70 --GGFLLGFDVAGIGLILTGAYLDIKALDKNAPKAAFKWTWCKMMLAGAVTMVTRLT 127
Db 67 SANEFYIAGFP--GITIIQT-VLEDVTSLSKLNPL-----LRSVPATDVYIFAVN 114
QY 128 EIIIPTEANSYNRKLNLSNIAF-----GGFEPSF 158
Db 115 EETTLGGFAHIYNGEIKRSF-IAYERVPEDNGIPGGFETPY 155

RESULT 15
US-09-860-670-134
; Sequence 134, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA137P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 134
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-670-134

Query Match 7.9%; Score 71.5; DB 9; Length 354;
Best Local Similarity 22.1%; Pred. No. 8.3;
Matches 36; Conservative 19; Mismatches 49; Indels 59; Gaps 5;
QY 1 MNKFLIIVLLAFVCFVSFAQADDSKSAFNIGAGEKLLAYETSKDPIVPELNL-FLGFG 59
Db 196 MGAFLAGVLLASSEYRHALES-----IEPFKGLLLGFFIGVG 234
QY 60 I-----GSFAQGDILGFLGFDVAGIGLITGAYLDIKALDKNAPKAAFKW----- 107
Db 235 MSIDFGTLLLENPLRIVILLGF-----LIKTAMLWLITARLPQVKNQRWFAVLIGQ 288
QY 108 -----TWGKMMLAGAVTMVTRLTEI 131
Db 289 SEFAFVVFGAAQMANVLEPEWAKSLTVALVSMATPILLVIL 331

Search completed: March 14, 2003, 14:37:05
Job time : 12 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Score				
1	737	81.2	179	2	B70104	hypothetical prote	
2	309	34.0	161	2	A70207	conserved hypothet	
3	301.5	33.2	170	2	F70243	conserved hypothet	
4	240	26.4	190	2	G70239	conserved hypothet	
5	93	10.2	779	2	AD1457	ATP dependent heli	
6	88.5	9.7	376	2	E55879	phosphate sugar ABC	
7	88.5	9.7	427	2	S55905	phosphotransferase	
8	88.5	9.7	451	2	C64976	galC protein - Esc	
9	88.5	9.7	451	2	G09390	PTS system galacti	
10	88.5	9.7	451	2	A85836	PTS system galacti	
11	87	9.6	1291	1	A48940	bontoxilysin (EC 3	
12	86.5	9.5	389	2	T28954	hypothetical prote	
13	86	9.5	272	2	H75011	hypothetical prote	
14	86	9.5	406	1	JC5041	fosmidomycin resis	
15	86	9.5	406	2	D06695	fosmidomycin resis	
16	86	9.5	406	2	H85545	fosmidomycin resis	
17	85	9.4	235	2	D63597	probable holocyto	
18	85	9.4	692	2	H70362	glycocon phosphory	
19	84.5	9.3	457	2	AH0898	PTS system, galact	
20	84	9.3	406	2	AH0563	fosmidomycin resis	
21	83.5	9.2	449	2	S73867	hypothetical prote	
22	82	9.0	357	2	H70346	undecaprenyl-phosp	
23	81.5	9.0	206	2	F64716	riboflavin synthas	
24	81.5	9.0	394	2	T13721	NADH2 dehydrogenas	
25	81.5	9.0	1185	2	AG2457	transcription repa	
26	81	8.9	174	2	TL1491	NADH2 dehydrogenas	
27	81	8.9	371	2	S55588	D-nopaline dehydro	
28	80.5	8.9	183	2	F72250	conserved hypothet	
29	80.5	8.9	315	2	H70203	xvlose operon regul	

A70207
conserved hypothetical protein BBA01 - Lyme disease spirochete plasmid A/lp54
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: A70207
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; H
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: A70207

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-161 <KLE>
A:Cross-references: GB:AE000790; NID:g2690224; PIDN:AAC66226.1; PID:g2690225; TIGR:BBA01
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 34.0%; Score 309; DB 2; Length 161;
Best Local Similarity 41.3%; Pred. No. 4.9e-21;
Matches 74; Conservative 25; Mismatches 60; Indels 20; Gaps 5;

QY 1 MNKFLIFILVIFCAFSSFAQDSSKSTFNILGAGEKFLVYETNNKDSLVPFLNLFGLFGIG 60

Db 1 MKKIFLILIFLGLTIQIFATKDTQNRKEGI-ESFNKYDKKKNPIGPFLLNLFPLFGIG 59

QY 61 SFAQGDILGSLILGFDVAGIGLILTGAYLDIKDFDNNAKKADFKWTWKGMMLAGV-VT 119

Db 60 SFVQGDYIGGSLVGNLGLAILWGTLNHRHRE-----TQLTGYILIGVGAS 107

QY 120 MAYTRLT-EIVLPFTFANNYNNRKLKNSLNIALGGFSPFDINMQASALGFLSFKKSY 177

Db 108 MVLTSYVSLIIPFTFANRHENLKKRLSRLAGFPNFDLGIN-----GFQLSFKKSY 161

RESULT 3

F70243

conserved hypothetical protein BBI31 - Lyme disease spirochete plasmid I/lp28-4
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: F70243

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70243

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-170 <KLE>

A:Cross-references: GB:AE000789; NID:g2690079; PIDN:AAC66194.1; PID:g2690091; TIGR:BBI31
A:Experimental source: strain B31

C:Genetics:
A:Genome: plasmid
C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 33.2%; Score 301.5; DB 2; Length 170;
Best Local Similarity 40.1%; Pred. No. 2.5e-20;
Matches 71; Conservative 23; Mismatches 70; Indels 13; Gaps 3;

QY 1 MNKFLIFILVIFCAFSSFAQDSSKSTFNILGAGEKFLVYETNNKDSLVPFLNLFGLFGIG 60

Db 7 MKKILILIFSLTIQIFATQD-KLEKSVGSETIMKYSEKATILAPLLNLFLLTIG 65

QY 61 SFAQGDILGSLILGFDVAGIGLILTGAYLDIKDFDNNAKKADFKWTWKGMMLAGVWTM 120

Db 66 SFVQGDYIGGAVLGSQGLGILCIAGNLTGHTDDTETRA-----TTGHIITIGVGII 118

QY 121 AVTRLTEIVLPFTFANNYNNRKLKNSLNIALGGFSPFDINMQASALGFLSFKKSY 177

Db 119 IASHIASLIIPFTFANKHNLKRLGIDTAGPEPNFDIGIS-----GFQLSFKKRY 170

RESULT 4

G70239

conserved hypothetical protein BBH41 - Lyme disease spirochete plasmid H/lp28-3
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: G70239

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: G70239
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-190 <KLE>

A:Cross-references: GB:AE000784; NID:g2690041; PIDN:AAC65989.1; PID:g2690045; TIGR:BE

A:Experimental source: strain B31

C:Genetics:
A:Genome: plasmid

C:Superfamily: Lyme disease spirochete plasmid hypothetical protein BBA01

Query Match 26.4%; Score 240; DB 2; Length 190;
Best Local Similarity 41.3%; Pred. No. 1.1e-14;
Matches 59; Conservative 22; Mismatches 50; Indels 12; Gaps 5;

QY 36 LVYETNNK-DSLVPFLNLFGLFGISFAQGDILGSLILGFDVAGIGLILTGAYLDIKD 94

Db 59 LQYEKQLIDLWIPVMLNLFPLPGVGSFVGQDYIGGCTLGFNLLGLTLTTGV-IQMK- 116

QY 95 FDNNAKKADFKWTWKGMMLAGVVTMAVTRLTEIVLPFTFANNYNNRKLKNSLNIALGGFE 154

Db 117 ---NLKKEPASIS-SMILLSGMLTFGSSYLSIYLPVLFEDRYKYNLMNRIIDELAGPE 172

QY 155 PSFDINMQASALGFLSFKKSY 177

Db 173 PNLDIGMN-----GFQLSFKKSY 190

RESULT 5

ADI457

ATP dependent helicase homolog lin0195 [imported] - Listeria innocua (strain Clp11261
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: ADI457

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.

A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.
A:Reference number: ABI077; MUID:21537279; PMID:11679669

A:Accession: ADI457

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-779 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC95428.1; PID:g16412614; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:
A:Gene: lin0195

Query Match 10.2%; Score 93; DB 2; Length 779;
Best Local Similarity 26.6%; Pred. No. 1-2;
Matches 45; Conservative 24; Mismatches 52; Indels 48; Gaps 9;

QY 14 AFSSFAQ-----DDSKSTFNILGAGEKFL-VYETNNKDSLVPFLNLFGLFGIGSFAQG 65

Db 614 APELFSEKYPEVRAQKQSQFMNEAREEFLLAAFTGRETFVGVFCV---LG---GVFSEG 667

QY 66 DILGSSILIGFDVAGIGLILTGAYLD- IKDFDNNAKKADFKWTWKGMMLA-----GVV 118

Db 668 VDLRGERLIGTAIVGGLAQMNVESDLIKDYNE-----TIGRGFDYAYIPGMKV 719

QY 119 TMAVTRLTE-----IVLPFTFAN-----NYNRKLKNSLNI 148

Db 720 LQAVGRVIRGKDGQGVLLIEERFAERYKALFPAHNNHARTVKNTLQI 768

RESULT 6
E95879
probable sugar ABC transporter permease protein Smb20315 [imported] - Sinorhizobium meliloti
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: E95879
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95879
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAC48701.1; PID:g15140174; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hynan, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
Rebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Smb20315
A:Genome: plasmid

Query Match 9.7%; Score 88.5; DB 2; Length 376;
Best Local Similarity 26.1%; Pred. No. 1.5;
Matches 49; Conservative 19; Mismatches 75; Indels 45; Gaps 10;

Qy 2 NKFLIFILVICAESSFAQDSSKSTFNLGAGEKFL-VYETNKKDSLVPLFLNLFGLFGIG 60
Db 65 NVLVVALAVF-----AGETFLSYNFQSMASQVPELALLALGVMLA 106
Qy 61 SFAQDILGSLILGFDVAVGIGLIL---TGAYLDIKDFDNNAKRAD--FKWTWKGMMMLA 115
Db 107 MIAGG-----GIDLSGIALANLAGVSYLLVRDW-VSADEAPLAFSLFAKALLI 157
Qy 116 GV-----VTMAYTRLEIVLFP---TFANNYRKLNSINIALGGPEPSFDINMGQAS 165
Db 158 GLAGLLNGALIAFAGLPPIIATLTGTLFTGLAVFTNGSAITLGYEP-LD-NFGNTP 215
Qy 166 ALGFGLSF 173
Db 216 VLGVPMCF 223

RESULT 7
S55905
phosphotransferase system enzyme II, galactitol specific, protein C - Escherichia coli
C:Species: Escherichia coli
A:Variety: strain EC3132
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 29-Sep-1999
C:Accession: S55905; S49083
R:Nobelmann, B.; Lengeler, J.W.
Biochim. Biophys. Acta 1262, 69-72, 1995
A:Title: Sequence of the gat operon for galactitol utilization from a wild-type strain B
A:Reference number: S55901; MUID:95290497; PMID:7772602
A:Accession: S55905
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-427 <NOB>
A:Cross-references: EMBL:X79837; NID:g599737; PIDN:CAA56230.1; PID:g508175
C:Genetics:
A:Gene: gatC
C:Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 9.7%; Score 88.5; DB 2; Length 427;
Best Local Similarity 27.5%; Pred. No. 1.7;
Matches 30; Conservative 15; Mismatches 33; Indels 31; Gaps 6;

Qy 46 LVPFLNLNF---LGFIGSFAQDILGSLILGFDVAVGIGLILTGAYLD-----IK 93
Db 17 MLPIVIIIFSKILG-----MKAGDCFKAGLHIGIGFVGIGLVI-GLMLDSIGPAAKAMAE 70
Qy 94 DFDNNAKADFKW-----TWCKGMMLAGV-----VTMAYTRLEIV 129
Db 71 NFDLNLHVVDVCGWPGSSPMTWASQIALVAIPAILNVNAMLTRTRVV 119
RESULT 8
C64976
gatC protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: C64976
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: C64976
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-451 <BLAT>
A:Cross-references: GB:AE000298; GB:U00096; NID:g1788395; PIDN:AAC75153.1; PID:g1788
A:Experimental source: strain K-12, substrain M61655
C:Genetics:
A:Gene: gatC
C:Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 9.7%; Score 88.5; DB 2; Length 451;
Best Local Similarity 27.5%; Pred. No. 1.8;
Matches 30; Conservative 15; Mismatches 33; Indels 31; Gaps 6;

Qy 46 LVPFLNLNF---LGFIGSFAQDILGSLILGFDVAVGIGLILTGAYLD-----IK 93
Db 17 MLPIVIIIFSKILG-----MKAGDCFKAGLHIGIGFVGIGLVI-GLMLDSIGPAAKAMAE 70
Qy 94 DFDNNAKADFKW-----TWCKGMMLAGV-----VTMAYTRLEIV 129
Db 71 NFDLNLHVVDVCGWPGSSPMTWASQIALVAIPAILNVNAMLTRTRVV 119
RESULT 9
G90990
PTS system galactitol-specific enzyme IIC [imported] - Escherichia coli (strain O157
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: G90990
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G90990
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-451 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA036318.1; PID:g13362364; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 050952
C:Genetics:
A:Gene: ECs2895
C:Superfamily: phosphotransferase enzyme II galactitol-specific

Query Match 9.7%; Score 88.5; DB 2; Length 451;
Best Local Similarity 27.5%; Pred. No. 1.8;
Matches 30; Conservative 15; Mismatches 33; Indels 31; Gaps 6;

Qy 46 LVPFLNLNF---LGFIGSFAQDILGSLILGFDVAVGIGLILTGAYLD-----IK 93
Db 17 MLPIVIIIFSKILG-----MKAGDCFKAGLHIGIGFVGIGLVI-GLMLDSIGPAAKAMAE 70

A; Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZ>
A; Cross-references: EMBL:Z11934; NID:g40383; PIDN:CAA77991.1; PID:g40384
R; Kuzano, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars
J. Biol. Chem. 267, 14721-14729, 1992
A; Note: Minimal essential domains specifying toxicity of the light chains of tetanus
A; Reference number: A42871; MUID:92340509; PMID:1634516
A; Accession: A42871
A; Status: nucleic acid sequence not shown
A; Molecule type: mRNA
A; Residues: 1-313, 'S', 315-451 <KUR>
A; Experimental source: strain Okra
A; Note: sequence extracted from NCBI backbone (NCBIP:109365)
R; Dasgupta, B.K.; Datta, A.
Biochimie 70, 811-817, 1988
A; Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity w
A; Reference number: S07155; MUID:89000987; PMID:3139097
A; Accession: S07155
A; Molecule type: protein
A; Residues: 2-29, 'M', 31-45 <DAS>
A; Accession: S08562
A; Molecule type: protein
A; Residues: 442-463, 'R', 465-467 <DA2>
R; Schmidt, J.J.; Sathiamoorthy, V.; Dasgupta, B.R.
Arch. Biochem. Biophys. 238, 544-548, 1985
A; Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
A; Reference number: S07128; MUID:85197963; PMID:3888113
A; Accession: S07128
A; Status: preliminary
A; Molecule type: protein
A; Residues: 2-16 <SCH1>
A; Accession: S08573
A; Status: preliminary
A; Molecule type: protein
A; Residues: 2-17 <SCH2>
A; Accession: S08574
A; Status: preliminary
A; Molecule type: protein
A; Residues: 442-459 <SCH3>
R; Schiavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta,
Nature 359, 832-835, 1992
A; Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by prote
A; Reference number: S27125; MUID:93063293; PMID:1331807
A; Contents: annotation
C; Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy
C; Genetics:
A; Gene: bont/b
C; Function:
A; Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C; Superfamily: tetanus toxin
C; Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F; 2-441/Product: bontoxilysin B light chain #status experimental <LGHT>
F; 442-1291/Product: bontoxilysin B heavy chain #status experimental <HVT>
F; 230,234/Binding site: zinc (His) #status predicted
F; 231/Active site: Glu #status predicted

	Query Match	9.6%; Score 87; DB 1; Length 1291;
	Best Local Similarity	22.8%; Pred. No. 7.5;
	Matches 38; Conservative 31; Mismatches 58; Indels 40; Gaps 7;	
Qy	37	VYETNKDS-----LVPEL--LNLFLGRIGSFAOG-DILGGSLILGFD---AV 79
	: : : :	
Db	602	VIEANKSNWDKTIADISLIYPYIIGLALVGNENAKGFENAFAGASILCEFIPELLIP 661
	: : : : : : : : : : : : : : :	
Qy	80	GIGLILTAYLD-----IKDFDNNAKADFKWTKGKMLAGVVTVMTATRL----- 125
	: : : : : : : : : : : : : : :	
Db	662	VWGAFLLESYIDKNKIITKDIALTRKNEKWSMDMYGLIVLAQLSTWTQFYTIKEGMVK 721
	: : : : : : : : : : : : : : :	
Qy	126	-----TEIVLPFTFANNYNKKLNLSINALIGGGPEPSPFINWGQA 164
	: : : : : : : : : : : : : : :	
Db	722	ALNYQAALEEEIIKYRY-NIYSEKESKNINIDFNINSKLNEGNIQA 767
	: : : : : : : : : : : : : : :	

RESULT 12

RESULT 12

T28954

hypothetical protein F28A12.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T28954

R:Sammons, L.; Murray, J.

submitted to the EMBL Data Library, July 1996

A:Description: The sequence of C. elegans cosmid F28A12.

A:Reference number: #20547

A:Accession: T28954

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-389 <SAM>

A:Cross-references: EMBL:U64851; PIDN:AAC47989.1; GSPDB:GN00023; CESP:F28A12.4

A:Experimental source: strain Bristol N2; clone F28A12

C:Genetics:

A:Gene: CESP:F28A12.4

A:Map position: 5

A:Introns: 104/2; 147/3; 175/3; 258/3; 301/3; 341/3

C:Superfamily: pepsin

Query Match 9.5%; Score 86.5; DB 2; Length 389;

Best Local Similarity 24.7%; Pred. No. 2.3;

Matches 40; Conservative 22; Mismatches 57; Indels 43; Gaps 8;

QY 22 DSKSTFNLGAGKEFLVYETNKKDSLVPFLLNLFGLGIGSFAOGDILGGSLILGFDVAVGI 81

DB 108 EKKRLFNKASSTYIA--TNR-----PWQI-----KYASGDAYG---TLGIDTWKI 148

QY 82 G-----LILTGAYLDIKDFDNNAKADFKWTWKGKMLAGVYTVMAVTRLTVEILVPTFA 135

DB 149 GSGEQAIPAIPRSYLGAVD---TVGDFKWSPKGIEGFLAFTALAVDNITPPII----- 199

QY 136 NNNKLNLSNIALGGPEPSFDINMGQASLGFLSPKSY 177

DB 200 -----NAINQGLLD-QPLFTWFGQAGPATSASGFTY 232

RESULT 13

H75011

hypothetical protein PAB1278 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000

C:Accession: H75011

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <KAW>

A:Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50534.1; PID:g545904

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1278

C:Superfamily: Pyrococcus abyssi hypothetical protein PAB1278

Query Match

Best Local Similarity 9.5%; Score 86; DB 2; Length 272;

Matches 37; Conservative 32; Mismatches 66; Indels 56; Gaps 6;

QY 3 KFLIFILVFCFAFFSFAODDSKSTENLGAGEKFLVYETNKK----- 43

DB 6 KALVIVIAIATLRFASIDTLMSL-LGLGLKLLGIRMSPNGFLFLFGLIALGLIMSDY 64

QY 44 ----DSLVPFLLNLFGLGIGSFAOQDILGGSLILGFDVAVGI-----GLILTGAYLDIK 93

DB 65 SVIVTSSIPFFLYSF----VWFQSNVLGNSVIAISMGEIVSSSTLGLIIVLLAFSFTV 120

QY 94 DFDNNAKADFKW-TWKGKMLAGVYTVMA-----VTRLTEIVLP 131

DB 121 DSYNSYEKLEVTWESYRFGLLISALVILTSIMFLALSRIDLSTQPVYKVVILLVLLVVP 180

QY 132 FTFANNYNRKL 142

DB 181 FLLVQEKKKV 191

RESULT 14

JC5041

fosmidomycin resistance protein - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002

C:Accession: JC5041; F64778

R:Fujisaki, S.; Ohnuma, S.; Horiuchi, T.; Takahashi, I.; Tsukui, S.; Nishimura, Y.;

Gene 175, 83-87, 1996

A:Title: Cloning of a gene from Escherichia coli that confers resistance to fosmidom

A:Reference number: JC5041; MUID:97074653; PMID:8917080

A:Accession: JC5041

A:Molecule type: DNA

A:Residues: 1-406 <FUJ>

A:Cross-references: DBJ:D73370; NID:g1019359; PIDN:BAAL1120.1; PID:g1019360

A:Experimental source: strain DH5alpha

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M

A.; Rose, D.J.; Mau, B.; Shaoh, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: F64778

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-406 <BLAT>

A:Cross-references: GB:AB000154; GB:U00096; NID:g1786683; PIDN:AAC73581.1; PID:g1786

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: fsr

C:Function:

A:Description: confers resistance to fosmidomycin

C:Superfamily: fosmidomycin resistance protein

C:Keywords: antibiotic resistance; transmembrane protein

F:104-120/Domain: transmembrane #status predicted <TM1>

F:181-197/Domain: transmembrane #status predicted <TM2>

F:225-241/Domain: transmembrane #status predicted <TM3>

F:267-283/Domain: transmembrane #status predicted <TM4>

F:294-310/Domain: transmembrane #status predicted <TM5>

F:320-336/Domain: transmembrane #status predicted <TM6>

F:358-374/Domain: transmembrane #status predicted <TM7>

F:381-397/Domain: transmembrane #status predicted <TM8>

Query Match 9.5%; Score 86; DB 1; Length 406;

Best Local Similarity 23.2%; Pred. No. 2.7;

Matches 39; Conservative 27; Mismatches 64; Indels 38; Gaps 6;

QY 6 IFILVIFCAFSSFAODDSKSTENLGAGEKFLVYETNKKDSLVPFLLNLFGLGIGSFAQG 65

DB 231 ILLILIFSKYFYMASSISYTYFL--MQKFGLSIQNALHLFAFLF-----AVAAG 279

QY 66 DILGSLILGFDVAVGIGLITGAYLDIKDFDNNAKADFKWTWKGKMLAGVYTVMAVTRL 125

DB 280 TWIGFV---GDKIGRKVYINGSLGVAPFTLILPYASLHWT-----GVLT----- 322

QY 126 TEIVLPFTFANNYNRKLNSLNIALGGPEPSFDINMGQASLGFLGSF 173

DB 323 --VIIGFTLASAFSAILVYAQLLPG-----RIGMVSGLFFGFAF 360

RESULT 15

D90695

fosmidomycin resistance protein [imported] - Escherichia coli (strain O157:H7, subst

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: D90695

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90695
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-406 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA03955.1; PID:g13359989; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs0532
C;Superfamily: fosmidmycin resistance protein

Query Match 9.5%; Score 86; DB 2; Length 406;
Best Local Similarity 23.2%; Pred. NO. 2.7;
Matches 39; Conservative 27; Mismatches 64; Indels 38; Gaps 6;
QY 6 IFILVIFCAFSSFAQDDSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFLGFGIGSFAQG 65
Db 231 ILLILIFSKYFYWASISSYTFYL--MQRFLSIQNAQLHLFAFLF-----AVAAG 279
QY 66 DILGGSLLILGFDVAGIGLITGAYLDIKDFDNNAKKADPKWTWKGMMLAGVVTMAVTRL 125
Db 280 TVIGGPV---GDKIGRKYVWGSILGVAPFTLILPYASLHWT-----GVLT----- 322
QY 126 TEIVLPFTFANNYNRKLKNSLNIALGGFEPFSDINMQASALGFLSF 173
Db 323 --VIIGFILASAFSAILVYAQLPG-----RIGMWSGLFFGFAP 360

Search completed: March 14, 2003, 14:25:44
Job time : 15.264 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:21:36 ; Search time 7.29214 Seconds
(without alignments)
1006.744 Million cell updates/sec

Title: US-09-508-487-23

Perfect score: 908

Sequence: 1 MNRFLIFILVIFCAFSSFAQ.....DINMGQASALGFLSPKKS 177

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.5	9.7	451	1 PTKC_ECOLI	P37189 escherichia
2	87	9.6	1290	1 BXB_CLOBO	P10844 clostridium
3	86	9.5	406	1 FSR_ECOLI	P52067 escherichia
4	85	9.4	235	1 CCDA_BACSU	P45706 bacillus su
5	85	9.4	692	1 PHSQ_AQUAE	O66932 aquifex aeo
6	83.5	9.2	449	1 YKE6_YEAST	P36091 saccharomyc
7	81	8.9	174	1 N06M_RABIT	O79438 oryctolagus
8	81	8.9	371	1 COXB_AGRIT4	Q914f5 pseudomonas
9	79.5	8.8	436	1 DCTA_PSEAE	P75596 escherichia
10	79.5	8.8	933	1 SIAP_CAMFE	P35827 campylobact
11	78.5	8.6	935	1 EAE_ECOLI	O31000 escherichia
12	78.5	8.6	936	1 EAE_CITFR	Q07591 citrobacter
13	78.5	8.6	939	1 EAE_ECO27	P19809 escherichia
14	78	8.6	264	1 YA96_MYCPN	P75596 escherichia
15	76.5	8.4	1103	1 VG37_BPARI	Q9u0b5 bacterioph
16	76.5	8.4	1938	1 MYHD_HUMAN	Q9ukx3 homo sapien
17	75.5	8.3	934	1 EAE_ECO57	P43261 escherichia
18	75.5	8.3	1077	1 HGPA_HAEIN	Q9za21 haemophilus
19	74.5	8.2	276	1 NOSY_PSEST	P19845 pseudomonas
20	74	8.1	359	1 OMPA_SERMA	P04845 serratia ma
21	74	8.1	362	1 FIBP_ADE08	P36845 human adeno
22	74	8.1	580	1 P69_MYCHR	P15362 mycoplasma
23	73.5	8.1	188	1 DGD_HELPRI	Q92kd0 helicobacte
24	73.5	8.1	419	1 CMLA_PSEAE	P32482 pseudomonas
25	72.5	8.0	836	1 VG26_BPML5	Q05233 mycobacteri
26	72	7.9	493	1 SECY_ARCFU	O28377 archaeoglob
27	72	7.9	607	1 LPPA_LAGIA	Q9cgl8 lactococcus
28	72	7.9	638	1 SG9_BOVIN	Q28039 bos taurus
29	72	7.9	670	1 OATP_HUMAN	P46721 homo sapien
30	72	7.9	1072	1 CARB_PSEAE	P38100 pseudomonas
31	71.5	7.9	392	1 FXB4_YEAST	Q06205 saccharomyc
32	71.5	7.9	444	1 DCTA_RHILE	Q01857 rhizobium l
33	71.5	7.9	853	1 PHS1_DICDI	Q00766 dictyosteli

RESULT 1

ID	PTKC_ECOLI	STANDARD;	PRT;	451 AA.
AC	P37189; P76411;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	PTS system, galactitol-specific IIC component (EIIC-GAT) (Galacticol-			
DE	permease IIC component) (Phosphotransferase enzyme II, C component).			
GN	GATC OR B2092 OR Z3255 OR ECS2895.			
OS	Escherichia coli, and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562, 83334;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=EC3132;			
RX	MEDLINE=95290497; PubMed=7772602;			
RA	Nobelmann B., Lengeler J.W.;			
RT	"Sequence of the gat operon for galactitol utilization from a			
RT	wild-type strain EC3132 of Escherichia coli.;"			
RL	Biochim. Biophys. Acta 1262:69-72(1995).			
[2]				
RP	SEQUENCE FROM N.A.			
RN	STRAIN=K12 / MG1655;			
RX	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.;"			
RL	Science 277:1453-1474(1997).			
[3]				
RP	SEQUENCE FROM N.A.			
RN	STRAIN=K12;			
RX	MEDLINE=97251358; PubMed=9097040;			
RA	Itoh T., Aiba H., Baba H., Fujita K., Hayashi K., Inada T.,			
RA	Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,			
RA	Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,			
RA	Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,			
RA	Saito N., Sampei G., Seki Y., Sivasubramanian S., Tagami H.,			
RA	Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;			
RT	"A 460-kb DNA sequence of the Escherichia coli K-12 genome			
RT	corresponding to the 40.1-50.0 min region on the linkage map.;"			
RL	DNA Res. 3:379-392(1996).			
[4]				
RP	SEQUENCE FROM N.A.			
RN	STRAIN=O157:H7 / ATCC 700927;			
RX	MEDLINE=9740935; PubMed=11206551;			
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,			
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,			
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Welch R.A., Blattner F.R.;			

P18387 capripoxvir
Q9piq8 campylobact
Q41144 ricinus com
P39457 penicillium
Q46100 campylobact
P26589 pneumonia v
P11818 human adeno
P41095 oryza sativ
Q9vaf0 drosophila
P72299 rhizobium m
P93563 solanum tub
Q986r8 rhizobium l

ALIGNMENTS

QY 126 TEIVLPPTFANNYRKLKNSLNIALGCFEPDFDINKQASALGFGLSF 173
 ::: | | : : : | : : : | | | : |
 Db 323 --VIIGFILASAFSAIVYAQELLPG-----RIGVWSGLFFGPAF 360

RESULT 4

CCDA_BACSU STANDARD; PRT; 235 AA.
 AC P45706;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome c-type biogenesis protein ccda.
 GN CCDA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97124194; PubMed=8969507;
 RA Rose M., Entian K.D.;
 RT "New genes in the 170 degrees region of the Bacillus subtilis genome
 RT encode DNA gyrase subunits, a thioresoxin, a xylanase and an amino
 RT acid transporter.";
 RL Microbiology 142:3097-3101(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97221596; PubMed=9068642;
 RA Schloett T., von Wachenfeldt C., Hederstedt L.;
 RT "Identification and characterization of the ccda gene, required for
 RT cytochrome c synthesis in Bacillus subtilis.";
 RL J. Bacteriol. 179:1962-1973(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bartorel M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinols S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Rapoport G., Prescott A.M.,
 RA Presecan E., Pujic P., Furnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Solido B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni K.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler H., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 RN [4]
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=168;

RX MEDLINE=20245543; PubMed=10781554;
 RA Schloett T., Hederstedt L.;
 RT "Efficient spore synthesis in Bacillus subtilis depends on the Ccda
 RT protein.";
 RL J. Bacteriol. 182:2845-2854(2000).
 CC -!- FUNCTION: REQUIRED FOR CYTOCHROME C SYNTHESIS AND STAGE V OF
 CC SPOULATION. MIGHT TRANSFER REDUCING EQUIVALENTS ACROSS THE
 CC CYTOPLASMIC MEMBRANE, PROMOTING EFFICIENT DISULFIDE BOND
 CC ISOMERIZATION OF PROTEINS LOCALIZED ON THE OUTER SURFACE OF THE
 CC MEMBRANE OR IN THE SPORE COAT.
 CC -!- PATHWAY: CYTOCHROME C SYNTHESIS SYSTEM II.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE DSD SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z73234; CAA97594.1; -;
 DR EMBL; X87845; CAA61116.1; -;
 DR EMBL; Z99113; CAA13677.1; -;
 DR Subtilist; BG11240; ccda.
 DR InterPro; IPR003834; Cytococh_TM.
 DR Pfam; PF02683; Dsbd; 1.
 KW Cytochrome c-type biogenesis; Sporulation; Transmembrane;
 KW Complete proteome.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 57 77 POTENTIAL.
 FT TRANSMEM 87 107 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 204 224 POTENTIAL.
 FT TRANSMEM 235 AA; 26007 MW; D94F3EG2EDD2AC29 CRC64;
 SQ
 Query Match 9.4%; Score 85; DB 1; Length 235;
 Best Local Similarity 24.4%; Pred. No. 0.95;
 Matches 48; Conservative 28; Mismatches 65; Indels 56; Gaps 9;
 QY 4 FLIFI-----LVIFCAFSSP-----AQDSSKSTFNLGAGERFLVYETNKKDSLVPFLN--- 52
 Db 14 FLFSFSPCLPLYPAPFLSVITGVSMDDVKT-----EKLLQKRSIFHTLC-FLIGFSV 65
 QY 53 LFLGFGIGSFAOGDI-----LGGSLILGFDVAGVIGLITLGTAYLDIKDFDNNAKKA 102
 Db 66 IFALGCTGTSFGISLFRDYHDAIRQIGALLIILFGITIGVFRPEAMKERRIHFHKPS 125
 QY 103 DF-----KWTGKGMMLAGVVTMAVTR-----LTEIVLPFTFANNYN 139
 Db 126 GFLGSLVIGMAFAAGWTPTCTGPIAAVITLAGTNPGSVPMYMLYVLGFAVPELLLSFFI 165
 QY 140 RKL-----KNSLNALGG 152
 Db 186 TKLWKIRKNQLFIMKAG 202
 RESULT 5
 PHSG_AQUAE
 ID PHSG_AQUAE STANDARD; PRT; 692 AA.
 AC O66932;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycogen phosphorylase (EC 2.4.1.1).
 GN GLGP OR AQ.717.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 OC Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]


```
DR EMBL: AJ001588; CAA04858.1; --
DR InterPro: IPR001457; Oxidored_q3.
DR Pfam: PF00499; oxidored_q3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 174 AA; 18728 MW; 18740BB6661D09E1 CRC64;

Query Match      8.9%; Score 81; DB 1; Length 174;
Best Local Similarity 23.1%; Pred. No. 1.6;
Matches 36; Conservative 24; Mismatches 54; Indels 42; Gaps 7;

QY 47 VPLLNL-----PLGF-----GIGSFAQGDILGSLILGFDVAGIG-----LILTGA 88
DB 4 VFLLSVMFVGVGVSSKPSPIYGLGLVSGGV-GCGIVLSFGGSFGLMFLIYLGG 62
QY 89 YLDIKDFDNNAKKADKFWTKGKMLAG-----VVTMAYRLTEIVLPTFEAN 136
DB 63 MLVVFGYTAMATEEPEWGSNMILGMFVLGVLMVWSDGVGVVDFPKMNG 122
QY 137 NYNRKLKNSLNIALGGFEPFSF--DINMQASALGFG 170
DB 123 DW-----VVFEGDEVGLIREDSMGVAALYSYG 149

RESULT 8
OOXB_AGRt4
ID OAXB_AGRt4 STANDARD; PRT; 371 AA.
AC Q59159;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Opine oxidase subunit B (EC 1.1.1.1) (Octopine oxidase subunit B).
GN OOXB.
OS Agrobacterium tumefaciens (strain Ach5).
OG Plasmid pTiAch5.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OC NCBI_TaxID=176298;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=94321320; PubMed=8045881;
RA Zanker H., Lurz G., Langridge U., Langridge P., Kreusch D.,
RA Schroeder J.;
RT 'Octopine and nopaline oxidases from Ti plasmids of Agrobacterium
RT tumefaciens: molecular analysis, relationship, and functional
RT characterization.';
RL J. Bacteriol. 176:4511-4517(1994).
CC -!- FUNCTION: OXIDATIVE CLEAVAGE OF OCTOPINE INTO L-ARGININE AND
CC PYRUVATE (BY SIMILARITY).
CC -!- PATHWAY: CATABOLIC UTILIZATION OF OCTOPINE.
CC -!- SUBUNIT: HETERODIMER OF A SUBUNIT AND A B SUBUNIT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: Z30328; CAA82987.1; --
DR InterPro: IPR000927; D_aa_oxidase.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01266; DAO; 1.
KW Oxidoreductase; Plasmid.
SQ SEQUENCE 371 AA; 39939 MW; 948705136822AC52 CRC64;

Query Match      8.9%; Score 81; DB 1; Length 371;
Best Local Similarity 27.3%; Pred. No. 3.5;
Matches 38; Conservative 18; Mismatches 43; Indels 40; Gaps 7;

QY 67 ILGSLILGFDVAGIG-----LILTGAYLDIKDFDNNAKADFK--WTWKGKMLLAGVV 118
DB 8 IIGGLVGIAGIANGIARSGTKPIVLGADLDLR-----ASRANFALVWVGKGLHAPHYA 62
```

```
QY 119 -----TMAYRLTEIVL-----PFTFANN-----YNRKLKNSLNIALGGFE 154
DB 63 LWSDSARRWPTWANTLLDSDGIDVGLQDGAFTFALSEELEARNQDMESIELETNGRA 122
QY 155 PSFDINMQAS----ALGFG 170
DB 123 PQFEVLDRQOTLDRVLGIG 141

RESULT 9
DCTA_PSEAE
ID DCTA_PSEAE STANDARD; PRT; 436 AA.
AC Q914F5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C4-dicarboxylate transport protein.
GN DCTA OR PA1183.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OC NCBI_TaxID=287;
RN [1]
SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Responsible for the transport of dicarboxylates such as
CC succinate, fumarate, and malate from the periplasm across the
CC inner membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SODIUM-DICARBOXYLATE SYMPORTER FAMILY
CC (SDF, TC 2.A.23).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: AE004548; AG04572.1; --
DR InterPro: IPR001991; Na/dico_symp.
DR Pfam: PF00375; SDF; 1.
DR PRINTS: PR00173; EDTNSPORT.
DR PROSITE: PS00713; NA_DICARBOXYL_SYMP.1; FALSE_NEG.
DR PROSITE: PS00714; NA_DICARBOXYL_SYMP.2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport;
KW Complete proteome.
FT TRANSMEM 9 28 POTENTIAL.
FT TRANSMEM 48 65 POTENTIAL.
FT TRANSMEM 78 100 POTENTIAL.
FT TRANSMEM 141 163 POTENTIAL.
FT TRANSMEM 189 211 POTENTIAL.
FT TRANSMEM 226 248 POTENTIAL.
FT TRANSMEM 290 312 POTENTIAL.
FT TRANSMEM 327 349 POTENTIAL.
FT TRANSMEM 354 376 POTENTIAL.
SQ SEQUENCE 436 AA; 45993 MW; DCA564DB31A5AC07 CRC64;

Query Match      8.8%; Score 79.5; DB 1; Length 436;
Best Local Similarity 37.1%; Pred. No. 5.6;
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:51 ; Search time 24.1966 seconds
(without alignments)
1507.249 Million cell updates/sec

Title: US-09-508-487-23
Perfect score: 908
Sequence: 1 MNKFLIFILVIFCAFSSFAQ.....DINMGASALGFLSFKKSY 177

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Watch 100%

Listing first 45 summaries

Database :

SPTREMBL21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	908	100.0	177	2 Q9X3V0	Q9x3v0 borrelia ga
2	786.5	86.6	178	2 Q9X309	Q9x3u9 borrelia af
3	737	81.2	179	16 O51065	O51065 borrelia bu
4	309	34.0	161	16 O50896	O50896 borrelia bu
5	301.5	33.2	170	16 O50885	O50885 borrelia bu
6	242.5	26.7	161	2 Q9S011	Q9s011 borrelia bu
7	240	26.4	190	16 O50696	O50696 borrelia bu
8	170.5	18.8	95	2 O31325	O31325 borrelia bu
9	95.5	10.5	289	16 Q9KXK6	Q9kxk6 streptomyce
10	93	10.2	779	16 Q92FB1	Q92fb1 listeria in
11	91	10.0	375	10 Q9LHX2	Q9lhx2 oryza sativ
12	90.5	10.0	427	2 Q93RE6	Q93re6 vibrio para
13	88.5	9.7	376	16 Q92WN1	Q92wn1 rhizobium m
14	86.5	9.5	389	5 Q22972	Q22972 caenorhabdi
15	86.5	9.5	427	2 Q8VS09	Q8vs09 klebsiella
16	86.5	9.5	447	16 Q8RFA1	Q8rfal fusobacteri

17	86	9.5	272	17 Q9UY78	Q9uy78 pyrococcus
18	86	9.5	406	16 Q8XD36	Q8xd36 escherichia
19	85.5	9.4	206	2 Q9L3R9	Q9l3r9 helicobacte
20	85.5	9.4	395	2 Q9XD03	Q9xd03 desulfitoba
21	84.5	9.3	457	16 Q8ZL08	Q8zlu8 salmonella
22	84.5	9.3	457	16 Q8ZJ33	Q8zj33 salmonella
23	84	9.3	406	16 Q8ZR98	Q8zr98 salmonella
24	84	9.3	406	16 Q8ZS88	Q8zs88 salmonella
25	83.5	9.2	399	2 Q9ANV6	Q9anv6 lactobacill
26	82	9.0	357	16 Q66804	Q66804 aquifex aeo
27	82	9.0	527	8 Q8SK12	Q8sk12 necator ame
28	81.5	9.0	206	2 O30550	O30550 helicobacte
29	81.5	9.0	206	16 O26094	O26094 helicobacte
30	81.5	9.0	246	16 Q8RQB2	Q8rgb2 fusobacteri
31	81.5	9.0	394	8 P92348	P92348 sobralia po
32	81.5	9.0	1185	16 Q8YLT1	Q8ylt1 anabaena sp
33	80.5	8.9	183	16 Q9XIG6	Q9xig6 thermotoga
34	80.5	8.9	315	16 O51771	O51771 borrelia bu
35	80.5	8.9	328	2 O54349	O54349 pasteurella
36	80.5	8.9	348	16 Q9CNN9	Q9cnn9 pasteurella
37	80	8.8	355	16 Q97EU1	Q97eu1 clostridium
38	80	8.8	589	16 Q93RS4	Q93rs4 streptomyce
39	80	8.8	626	16 O51504	O51504 borrelia bu
40	80	8.8	690	8 Q8WJ22	Q8wj22 brachypodiu
41	80	8.8	1291	2 Q80777	Q80777 clostridium
42	79.5	8.8	226	2 Q99559	Q99559 mycoplasma
43	79.5	8.8	1651	2 Q9F0J1	Q9f0j1 streptomyce
44	79	8.7	325	16 Q8REW2	Q8rew2 fusobacteri
45	79	8.7	414	16 Q92IV6	Q92iv6 rickettsia

ALIGNMENTS

RESULT 1

Q9X3V0 PRELIMINARY; PRT; 177 AA.
ID Q9X3V0
AC Q9X3V0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Membrane protein Pl3.
GN Pl3.
OS Borrelia garinii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP90;
RX MEDLINE=21189251; PubMed=11292755;
RA Noppa L., Ostberg Y., Lavrinovicha M., Bergstrom S.;
RT "Pl3, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains.";
RL Infect. Immun. 69:3323-3334(2001).
DR EMBL; AF085741; AAD28362.1; .
SQ SEQUENCE 177 AA; 19308 MW; 429928CC426E1111 CRC64;

Query Match 100.0%; Score 908; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.6e-73;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKFLIFILVIFCAFSSFAODSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFLGFGIG 60

Db 1 MNKFLIFILVIFCAFSSFAODSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFLGFGIG 60

QY 61 SFAQGDILGSLILGFDVAVGILGILITGAYLDIKDFDNNAKKADFKWTGKGMVLGVVTM 120

Db 61 SFAQGDILGSLILGFDVAVGILGILITGAYLDIKDFDNNAKKADFKWTGKGMVLGVVTM 120

QY 121 AVTRLTEIVLPFTFANNYNRKLKNSLIALGGFPEPSFDINNGQASALGFLSFKKSY 177

Db 121 AVTRLTEIVLPFTFANNYNRKLKNSLIALGGFPEPSFDINNGQASALGFLSFKKSY 177


```
RESULT 5
O50885 ID O50885 PRELIMINARY; PRT; 170 AA.
DR EMBL; AE001584; AAF07707.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 161 AA; 14D4906CC8107CD2 CRC64;

Query Match 26.7%; Score 242.5; DB 2; Length 161;
Best Local Similarity 39.3%; Pred. No. 3.9e-14;
Matches 59; Conservative 25; Mismatches 49; Indels 17; Gaps 6;

QY 18 FAQDSDSKSTFNLAGAE--KELVYETNKDSLVPFLNLFGLFGIGSFAQGDILGGSLLIG 75
Db 4 FAQEKLEK----GVGDIATVMKYESKATILAPLLNLLSLGIGSFVQGDYIGGCALLG 59
QY 76 FDAYGIGLILTCAYL--DIKDFDNNAKKADFKWTGKGMMLAGV--VTMAYTRTEIVLP 131
Db 60 SOVLG-GILIMAGYMTGDFVTESTA-----TVITGVLSGIGLTIASAAYITGIIP 112
QY 132 FTFANNYNRKLNLSNIALGGFEPFSFDNM 161
Db 113 KFFANNYNADLKKRLGIALAGLEPNFDIGI 142

RESULT 7
O50696 ID O50696 PRELIMINARY; PRT; 190 AA.
AC O50696;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBH41.
GN BBH41.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RA "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL; AE000789; AAC66194.1; -.
DR TIGR; BB131; -.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 170 AA; 18227 MW; 0E71B633EA80FB22 CRC64;

Query Match 33.2%; Score 301.5; DB 16; Length 170;
Best Local Similarity 40.1%; Pred. No. 2.2e-19;
Matches 71; Conservative 23; Mismatches 70; Indels 13; Gaps 3;

QY 1 MNKFLIFLTVFCFSSPAODSKSTFNLAGAEKFLVYETNKDSLVPFLNLFGLFGIG 60
Db 7 MKKILTLILFSLTIQIPATOD-KLEKSVGSIETIMKYSEKATILAPFLNLFGLFGIG 65
QY 61 SFAQGDILGSLILGFDVAVGIGLITLTCAYLTKDIFDNNAKKADFKWTGKGMMLAGVVTM 120
Db 66 SFVQGDYIGGAVLGSQSLGILCTAGNLTGTDDETRA-----TTGHIITIGVCTI 118
QY 121 AVTRLTEIVLFTFANNYNRKLNLSNIALGGFEPFSFDINMQASALGFGLSFKSY 177
Db 119 IASHIASLIIPFTFANKHNANLKKRLGIDIAGFEPNFDIGIS-----GFQLSFKRY 170

RESULT 6
Q9S011 ID Q9S011 PRELIMINARY; PRT; 161 AA.
AC Q9S011;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Conserved hypothetical protein.
GN BQ006.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,
RA White O., Dodson R., Hickey E.K., Gwinn M., Peterson J., van-Vugt R.,
RA Palmer N., Haft D., Rosa P., Stevenson B.;
RT "A bacterial genome in flux: The twelve linear and nine circular
RT extrachromosomal DNAs in an infectious isolate of the Lyme disease
RT spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 0:0-0(1999).
```

```
ID O31325 PRELIMINARY; PRT; 95 AA.
AC O31325;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (clone 8) s3.
GN S3.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N40;
RX MEDLINE=96305348; PubMed=8688460;
RA Feng S., Das S., Barthold S.W., Fikrig E.;
RT "Characterization of two genes, p11 and p5, on the Borrelia
RL Blochm. Biophys. Acta 1307:270-272(1996).
EMBL; L41151; AAB62281.1; -.
SQ SEQUENCE 95 AA; 10739 MW; D12FC791F8920C3A CRC64;

Query Match 18.8%; Score 170.5; DB 2; Length 95;
Best Local Similarity 45.0%; Pred. No. 5.6e-08;
Matches 36; Conservative 11; Mismatches 32; Indels 1; Gaps 1;

QY 1 MKKFLIFILVFCFAFSFAODSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIG 60
DB 1 MKKFTLLIFLGLTIEIFATKQNTKEGI-ESFNKYDKEKNPGIPFLNLFGLFGIG 59

QY 61 SFAQGDILGSLILGFDVANG 80
DB 60 SFVQGDYIGGSGVLGNLLG 79

RESULT 9
Q9XXK6 PRELIMINARY; PRT; 289 AA.
ID Q9XXK6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SCO2214 OR SCC53.05.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 413:141-147(2002).
EMBL; AL357591; CAB93434.1; -.
SQ SEQUENCE 289 AA; 29123 MW; 89AD68D9D805BE10 CRC64;

Query Match 10.5%; Score 95.5; DB 16; Length 289;
Best Local Similarity 28.8%; Pred. No. 1;
Matches 36; Conservative 13; Mismatches 43; Indels 33; Gaps 5;

QY 71 SLILGFDVANGIG---LILTGAYLIDIKDFDNNAKADFKTWKGKMLAGVVTWAVTRLTE 127
DB 5 SLTRG-DGVVIGRAVLLIIASFLDIYSIDGAPDSADIPSLWGSVVLGVLAGVGAAL 63
```

```
QY 128 IVLPFTFANNYNNRKLKNSLNIA---LGGFEFSDI-----NMGOASA 166
DB 64 VVV-----NRAMPQPKVAGLDLQFGIATVFAAWCALGNIFDPAGAFDNDGGASG 115

QY 167 LGFGL 171
DB 116 IGAGM 120

RESULT 10
Q92FB1 PRELIMINARY; PRT; 779 AA.
ID Q92FB1;
AC Q92FB1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein lin0195.
GN Lin0195.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
EMBL; AL596163; CAC95428.1; -.
DR ListiList; LIN00195; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 779 AA; 89623 MW; AB7B7D066C715261 CRC64;

Query Match 10.2%; Score 93; DB 16; Length 779;
Best Local Similarity 26.6%; Pred. No. 5.7;
Matches 45; Conservative 24; Mismatches 52; Indels 48; Gaps 9;

QY 14 AFSSFAQ-----DDSKSTFNLGAGEKFL-VYETNKKDSLVPFLNLFGLFGIGSFAQG 65
DB 614 AFELSEKYPEVRAQKQSPFMNEAREEFLAAFETGRETFVGFV---LG---GVFSEG 667

QY 66 DILGSSILIGFDVANGIGLILTGAYLD-IKDFDNNAKADFKTWKGKMLA-----GVV 118
DB 668 VDLRGERLIGTAIVGGLAQMNVESDLIKDYNE-----TIGRGFDVAYIPGMKV 719

QY 119 TMAVTRLTE-----IVLPFTFAN-----NYYRKLKNSLNI 148
DB 720 LQAVGRVIRCEKDGQGVVLLIEERFATERYKALFPAHNNHARTVKNLTQI 768

RESULT 11
Q9LHX2 PRELIMINARY; PRT; 375 AA.
ID Q9LHX2;
AC Q9LHX2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to Arabidopsis thaliana chromosome II sequence from clones
DE T22013 (OSUNBa0036717.30 protein).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
```

```
OC Ehrhartoidea; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
  clone:P0515G01.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
  clone:OSJNBa0036J17.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP001633; BAA94232.1; -.
DR EMBL; AP003104; BAB55738.1; -.
DR InterPro; IPR001087; Lipase_GDSL.
SQ SEQUENCE 375 AA; 40234 MW; 238D34ACAD934E93 CRC64;

Query Match 10.08; Score 91; DB 10; Length 375;
Best Local Similarity 24.38; Pred. No. 3.6;
Matches 42; Conservative 26; Mismatches 61; Indels 44; Gaps 8;

Qy 1 MNKEL-IFILVIFCAFSSFAODDSKSTFNLG-----AGEKFLVYETNKKDSLVS---PFL 51
D 1 MOKVLIQIIVLLSFSFSEVETDVASIFSGDSFDGTGNIVLIYGPARTDLVMKPPYGM 60
Qy 52 NLFGLFGTGAQGDILGGSLILGFDVAVIGL-ILTGAYLDIKDFDNNAKKADFKWTWGK 110
D 61 TFF-----DHPGRLSDGLIIDFIAEALGLPLLPSPFAANRSFEHCA----- 103
Qy 111 GMLAGVVTMAVTRTEIVLPFTFANNYNRKLNLSNLALG-----GFEP 156
D 104 -----NFATAGCTALDRAFFVANNFT--VMSPFNISLQDLGWLGDGMRPS 146

RESULT 12
Q93RE6 93RE6 PRELIMINARY; PRT; 427 AA.
AC Q93RE6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Adhesin.
GN VAG.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RA Akeda Y., Honda T.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047560; BAB59008.1; -.
DR InterPro; IPR02528; Mate.
DR Pfam; PF01554; UPF0013.2.
DR TIGRfams; TIGR00797; mate; 1.
SQ SEQUENCE 427 AA; 45909 MW; 0B608B9DC70E3C71 CRC64;

Query Match 10.08; Score 90.5; DB 2; Length 427;
Best Local Similarity 26.18; Pred. No. 4.6;
Matches 52; Conservative 23; Mismatches 65; Indels 59; Gaps 11;

Qy 5 LIFILVIFCAFSSFAQ-----DDSKSTFNLGAGEKFLVYETNKKDSLVS----- 47
D 73 MVFAALAVCLFVFPQPTGLMTDSQEVIELGS--SYLVISSASMFACVAVSNVGLRA 130
Qy 48 ---PFLNLFGLFGIGS--FAQGDILGSLILGFDVAVG-----TGLITGA-----Y 89
D 131 MHQFGLSTFFSGIGLSNVLNWLIFGH--LGFPALGITGAATATVISGAIEVGCLFGY 188

Ehrhartoidea; Oryzae; Oryza.
NCBI_TaxID=4530;
SEQUENCE FROM N.A.
STRAIN=CV. NIPPONBARE;
Sasaki T., Matsumoto T., Yamamoto K.;
"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0515G01.";
Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
SEQUENCE FROM N.A.
STRAIN=CV. NIPPONBARE;
Sasaki T., Matsumoto T., Yamamoto K.;
"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OSJNBa0036J17.";
Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AP001633; BAA94232.1; -.
EMBL; AP003104; BAB55738.1; -.
InterPro; IPR001087; Lipase_GDSL.
SEQUENCE 375 AA; 40234 MW; 238D34ACAD934E93 CRC64;

Query Match 9.78; Score 88.5; DB 16; Length 376;
Best Local Similarity 26.18; Pred. No. 6;
Matches 49; Conservative 19; Mismatches 75; Indels 45; Gaps 10;

Qy 2 NKFLIFILVIFCAFSSFAQDDSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIG 60
D 65 NVLVVALAVF-----AGETFLSLNFQSMSAQVPELALLALGVMLA 106
Qy 61 SFAQGDILGSLILGFDVAVIGLIL---TGAYLDIKDFDNNAKKAD--FKWTWGKMLA 115
D 107 MIAGGG-----GIDLSGIALANLAGVGYLLVRDW-VSADEAPLAFSWLFAAMALLI 157
Qy 116 GV-----VTMAVTRTEIVLPF---TFANNYNRKLNLSNLALGGEPEPFDINMGAS 165
D 158 GLAGLLNGALIAFAGLPIIATIGTOLFTGTLAVAFNGSAITLGYIEP-LD-NFGWTP 215
Qy 166 ALGFGLSF 173
D 216 VLGVPMCF 223

RESULT 14
Q22972 22972 PRELIMINARY; PRT; 389 AA.
AC Q22972;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE F28A12.4 protein.
GN F28A12.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
  Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lighning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Snaledon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wooldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Sammons, L, Murray J.;
RT "The sequence of C. elegans cosmid F28A12."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64851; AAC47989.1;
DR HSP; P00797; 2REN
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; P00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN1.
SQ SEQUENCE 389 AA; 42238 MW; A76270702C900C98 CRC64;

Query Match          9.5%; Score 86.5; DB 5; Length 389;
Best Local Similarity 24.7%; Pred. No. 9.5;
Matches 40; Conservative 22; Mismatches 57; Indels 43; Gaps 8;

QY 22 DSKSTFNLAGKFLVYETNKKDSLVPFLNLFGLFGIGSFAQGDILGSLILGFDVAVGI 81
Db 108 EKRLFNKASSYIA--TNR-----PWQI-----KYASGDVYG---TLGIDTFVKI 148

QY 82 G-----LIITGAYLDIKFDNNNAKKADFKWTGKGMMLAGVVTMAYTRLTEIVLPPTFA 135
Db 149 GGSGEAGLAIPRSLVGVAD---TVGSDFKWSPKEGIFGLAFTALAVDNTITPPII----- 199

QY 136 NNYNRKLKNSLNTALGFPFSPFDINMGQASALGFLSFKSY 177
Db 200 -----NAINOGLLD-QPLETTWFQGGAPGTSASGAFTY 232

RESULT 15
QYVS09 PRELIMINARY; PRT; 427 AA.
AC QYVS09;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GATC.
GN Klebsiella oxytoca.
OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=571;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-M5A1;
RA Shakeri-Garakani A., Brinkkoetter A., Schmid K., Turgut S.,
RA Lengeler J.W.;
RT "The genes and enzymes for the catabolism of galactitol, D-tagatose,
RT and related carbohydrates in Klebsiella oxytoca M5a1 and other enteric
RT bacteria reflect a modular evolution."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF416702; AAL60172.1;
DR InterPro; IPR004703; Gal_spec.IIC.
DR InterPro; IPR001865; Ribosomal_S2.
```

```
DR Pfam; PF03611; EIIC-GAT; 1.
DR TIGRFAMS; TIGR00827; EIIC-GAT; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2.1; UNKNOWN1.
SQ SEQUENCE 427 AA; 45406 MW; 1D1817F0287F7F56 CRC64;

Query Match          9.5%; Score 86.5; DB 2; Length 427;
Best Local Similarity 25.4%; Pred. No. 11;
Matches 36; Conservative 18; Mismatches 39; Indels 49; Gaps 8;

QY 46 LVPFLLNLF---LGFGIGSFAQGDILGSLILGFDVAVGIGLITGAYLD-----IK 93
Db 17 MLPVIVIIIFSKILGMKL-----GDFKSGLHIGIGFVGIGLVI-GLMLDSIGPAKAAAE 70

QY 94 DFDNNNAKKADFKW-----TWGKGMMLAGV-----VTMAVTRLTEIV-----LP 131
Db 71 HFQLNLHVIDIGWPGSSPMTWASQIALVAIPATIAVNIEMLVTRVTVVNDIWNHMT 130

QY 132 FTFANNYNRKLKNSLNTALGGF 153
Db 131 FTGA-----MLHIATGSY 143
```

Search completed: March 14, 2003, 14:27:07
Job time : 27.1966 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 14, 2003, 14:23:06 ; Search time 10.9382 Seconds
(without alignments)
476.116 Million cell updates/sec

Title: US-09-508-487-23
Perfect score: 908
Sequence: 1 MNKFLIPILVIFCAFSSFAQ.....DINMGASALGFLGSLFKKSY 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	87	9.6	858	4	US-09-255-829-22
2	87	9.6	858	4	US-09-255-829-29
3	87	9.6	1169	4	US-09-255-829-20
4	77	8.5	243	4	US-09-134-001C-2961
5	74.5	8.2	934	4	US-08-840-466A-19
6	74.5	8.2	934	4	US-09-696-188B-19
7	71	7.8	612	3	US-09-295-186-16
8	70.5	7.8	95	4	US-08-851-362D-37
9	70.5	7.8	116	4	US-08-851-362D-49
10	70.5	7.8	344	4	US-09-134-001C-4175
11	70.5	7.8	581	4	US-09-393-627B-1
12	69	7.6	376	1	US-08-608-241-2
13	69	7.6	376	2	US-08-922-182-2
14	69	7.6	376	2	US-09-919-953-2
15	69	7.6	376	4	US-09-192-983-2
16	69	7.6	415	4	US-09-134-001C-5101
17	68	7.5	165	4	US-08-858-207A-520
18	68	7.5	539	2	US-08-467-963C-2
19	68	7.5	539	2	US-08-838-189D-2
20	68	7.5	539	3	US-08-852-344D-2
21	68	7.5	539	3	US-08-344-639E-2
22	68	7.5	539	4	US-08-467-969A-2
23	68	7.5	539	4	US-08-467-961A-2
24	68	7.5	539	4	US-08-001-554A-2
25	68	7.5	691	4	US-09-620-412C-313
26	67.5	7.4	349	4	US-09-134-001C-4004
27	67.5	7.4	538	4	US-09-134-001C-4633

28	67.5	7.4	1861	2	US-08-790-912-4	Sequence 4, Appl
29	67	7.4	239	2	US-08-933-750C-22	Sequence 22, Appl
30	67	7.4	239	4	US-09-234-613-22	Sequence 22, Appl
31	67	7.4	337	1	US-08-312-387B-5	Sequence 5, Appl
32	67	7.4	337	1	US-08-312-387B-12	Sequence 12, Appl
33	67	7.4	337	1	US-08-683-426-5	Sequence 5, Appl
34	67	7.4	337	1	US-08-683-426-12	Sequence 12, Appl
35	67	7.4	337	1	US-08-683-458-5	Sequence 5, Appl
36	67	7.4	337	1	US-08-683-458-12	Sequence 12, Appl
37	67	7.4	337	2	US-08-878-360-5	Sequence 5, Appl
38	67	7.4	337	2	US-08-878-360-12	Sequence 12, Appl
39	67	7.4	337	3	US-08-478-140B-5	Sequence 5, Appl
40	67	7.4	337	4	US-09-333-412-5	Sequence 5, Appl
41	67	7.4	337	4	US-09-333-412-12	Sequence 12, Appl
42	67	7.4	337	4	US-09-338-943-5	Sequence 5, Appl
43	67	7.4	346	4	US-09-134-001C-5196	Sequence 5196, Ap
44	67	7.4	412	4	US-09-423-439-10	Sequence 10, Appl
45	67	7.4	639	4	US-09-134-001C-5661	Sequence 5661, Ap

ALIGNMENTS

RESULT 1
US-09-255-829-22
; Sequence 22, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255.829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782.893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 858 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-255-829-22

Query Match 9.6%; Score 87; DB 4; Length 858;
Best Local Similarity 22.8%; Pred. NO. 0.22;
Matches 38; Conservative 31; Mismatches 59; Indels 40; Gaps 7;

```

Db      722 ALNYQAQALEIHKYRY-NIYSEKEKSNINIDFNDSKLNKNEGINOA 767
RESULT 3
US-09-255-829-20
; Sequence 20, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNB, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782,893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1169 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-255-829-20

Query Match          9.6%; Score 87; DB 4; Length 11
Best Local Similarity 22.8%; Pred. No. 0,33;
Matches 39; Conservative 31; Mismatches 58; Indels

QY      37 VYETNKKDS-----LVDFPL---LNFLEFGFGISFPAQG-DILGGSLIL
      |||||:::|||||:::|||||:::|||||:::|||||:::
Db      602 VIEAKNNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFETIAGASIL
      |||||:::|||||:::|||||:::|||||:::|||||:::
QY      80 GIGLILTCAYLD-----IKDFDNNAKADFKWTGKGMGLAGVVTNATVTEL
      |||||:::|||||:::|||||:::|||||:::|||||:::
Db      662 VVGAFLLSEYIDNKNKIILKTIDNALTNRKEKWSMDYGLIVAQWLSTVNTQF
      |||||:::|||||:::|||||:::|||||:::|||||:::
QY      126 -----TEIVLPPTFFANNYNRKLKNSLNIALLGGFEPESFDINMGQA 164
      |||||:::|||||:::|||||:::|||||:::|||||:::
Db      722 ALNYQAQALEIHKYRY-NIYSEKEKSNINIDFNDSKLNKNEGINOA 767

RESULT 4
US-09-134-001C-2961
; Sequence 2961, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

```

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2961
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2961

Query Match 8.5%; Score 77; DB 4; Length 243;
Best Local Similarity 30.0%; Pred. No. 0.94;
Matches 36; Conservative 14; Mismatches 54; Indels 16; Gaps 5;

Qy 58 GIGSFAQGDILGSLILGFDVAGLILTG-AYLDIKDFDNNAKKADFKWTWKGMMLAG 116
Db 16 GINSFSLFILTAGWAAFPNCGIALLPSIISYL-----IGETKDSFRVAIFKGLGLGG 71
Qy 117 VVTMAVTRLTETVLVPTFFANNYRNKLNIAALGGPEPSFDINMQASA-LGFGLSFKK 175
Db 72 AMT-----TCFLTIFVLG-----LLIGGLGSALTGFIPILSLVMGILIALLLGLMFLGK 121

RESULT 5
US-08-840-466A-19
; Sequence 19, Application US/08840466A
; Patent No. 6261561
; GENERAL INFORMATION:
; APPLICANT: Stewart, C. Neal
; McKee, Marian L.
; O'Brien, Alison D.
; Wachtel, Marian R.
; TITLE OF INVENTION: Method Of Stimulating An Immune Response
; By Administration Of Host Organisms That Express Intimin
; Alone Or As A Fusion Protein With One Or More Other
; Antigens.
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,466A
; FILING DATE: 18-Apr-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Boone, Laural S.
; REGISTRATION NUMBER: 43,505
; REFERENCE/DOCKET NUMBER: 04995.0029-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-840-466A-19

Query Match 8.2%; Score 74.5; DB 4; Length 934;
Best Local Similarity 21.7%; Pred. No. 6.7;
Matches 34; Conservative 22; Mismatches 46; Indels 55; Gaps 7;

Qy 22 DSKSTENLGAGEKFLVYETNKKDSLVPFLNLFLGFGIGSFAQGDILGSLILGFDVAGI 81
Db 252 DSRTANLGAGQRF-----FLPANMLGYNV--FIDQFSGDNTRLG----- 290
Qy 82 GLILTGAYLDIKDFDNNAKKADFKW-TWKGMMLAGVVTMAVTRLTETVLVPTFFANNYRN 140
Db 291 ---IGGEYW--RDYFKSVNGYFRMRWRHESY-----PALGAKLIYEQYY 354

Qy 141 KLKNSLNIALGGPEPSFDINMQASALGFLSPKKS 177
Db 325 RPANGFDIRFNGYLPST-----PALGAKLIYEQYY 354

RESULT 6
US-09-696-188B-19
; Sequence 19, Application US/09696188B
; Patent No. 6406885
; GENERAL INFORMATION:
; APPLICANT: Stewart, C. Neal
; McKee, Marian L.
; O'Brien, Alison D.
; Wachtel, Marian R.
; TITLE OF INVENTION: Method Of Stimulating An Immune Response
; By Administration Of Host Organisms That Express Intimi
; Alone Or As A Fusion Protein With One Or More Other
; Antigens.
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/696,188B
; FILING DATE: 26-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/840,466
; FILING DATE: 1997-04-18
; ATTORNEY/AGENT INFORMATION:
; NAME: Boone, Laural S.
; REGISTRATION NUMBER: 43,505
; REFERENCE/DOCKET NUMBER: 04995.0029-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-696-188B-19

Query Match 8.2%; Score 74.5; DB 4; Length 934;
Best Local Similarity 21.7%; Pred. No. 6.7;

```
Matches 34; Conservative 22; Mismatches 46; Indels 55; Gaps 7;

QY 22 DSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAQGDILGGSLILGFDVAGI 81
||:||||||:|
Db 252 DSREFANLAGQRE-----FLPANMLGYNV--FIDQDFSGDNTRLG----- 290
||:||||||:|

QY 82 GLIITGAYLDIKDDNNAKADFKW-TWVGKMLAGVVTMAVTRLTEIVLPFTFANNYNR 140
||:||||||:|
Db 291 ---IGGEYW--RDYFKSVNGVYFMRRWHESY-----HKKDYDE 324
||:||||||:|

QY 141 KLNLSNIALGGFPEFSDINNGOASALGFLGSLFKSY 177
||:||||||:|
Db 325 RPANGFDIRENGYLPYSY-----PALCAKLIYEQY 354
||:||||||:|

RESULT 7
US-09-295-186-16
; Sequence 16, Application US/09295186B
; Patent No. 6127137
; GENERAL INFORMATION:
; APPLICANT: Hasida, Miyoko
; APPLICANT: Tsutsumi, No. 6127137iko
; APPLICANT: Halkier, Torben
; APPLICANT: Stringer, Mary Ann
; TITLE OF INVENTION: An Acidic Phospholipase, Production, and
; FILE REFERENCE: Methods of Using Thereof (As Amended)
; CURRENT APPLICATION NUMBER: US/09/295,186B
; CURRENT FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 1215/96
; PRIOR FILING DATE: 1996-10-31
; PRIOR APPLICATION NUMBER: PCT/DK97/00490
; PRIOR FILING DATE: 1997-10-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Penicillium notatum
US-09-295-186-16

Query Match 7.8%; Score 71; DB 3; Length 612;
Best Local Similarity 26.4%; Pred. No. 9.4;
Matches 33; Conservative 12; Mismatches 38; Indels 42; Gaps 5;

QY 14 AFSSFAQDSSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFGLFGIGSFAQGDI----- 67
||:||||||:|
Db 109 AFDS-RTDNATATGHLGLLQSATY-----ISGLSGGSLGLGIYINFT 152
||:||||||:|

QY 68 -----LGGSLILGFDVAGIGLI-LTGAYLDIKDDNNAKADFKWT----- 107
||:||||||:|
Db 153 TVDKLQTHEAGSVWQFGNSIIIEGPDAGGIQLLDSAGYKDLADVDGKKKAGFDTTLTDI 212
||:||||||:|

QY 108 WKGKM 112
||:|
Db 213 WGRAL 217

RESULT 8
US-08-851-362D-37
; Sequence 37, Application US/08851362D
; Patent No. 6235883
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Ava
; APPLICANT: Yang, Xiao-Dong
; APPLICANT: Gallo, Michael
; APPLICANT: Jia, Xiao-Chi
; TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal
; FILE REFERENCE: Growth Factor Receptor
; CURRENT APPLICATION NUMBER: US/08/851,362D
; CURRENT FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 49
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 95
; TYPE: PRT
; ORGANISM: human
US-08-851-362D-37

Query Match 7.8%; Score 70.5; DB 4; Length 95;
Best Local Similarity 23.0%; Pred. No. 0.83;
Matches 17; Conservative 13; Mismatches 27; Indels 17; Gaps 2;

QY 98 NAKKADFKWTW-----GKGMMLAGVVTMAVTRLTEIVLPFTFANNYNRKLKNSLNIALGG 152
:||||||:|
Db 5 SVSSGDYWTWIRQSPGKGLEWIGHY-----YSGNTNPNPSLKSLRTISIDT 52
:||||||:|

QY 153 FEPSFDINMGQASA 166
:||||||:|
Db 53 SKTQFSLKLSVTA 66
:||||||:|

RESULT 9
US-08-851-362D-49
; Sequence 49, Application US/08851362D
; Patent No. 6235883
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Ava
; APPLICANT: Yang, Xiao-Dong
; APPLICANT: Gallo, Michael
; APPLICANT: Jia, Xiao-Chi
; TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal
; FILE REFERENCE: Growth Factor Receptor
; CURRENT APPLICATION NUMBER: US/08/851,362D
; CURRENT FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 116
; TYPE: PRT
; ORGANISM: human
US-08-851-362D-49

Query Match 7.8%; Score 70.5; DB 4; Length 116;
Best Local Similarity 23.0%; Pred. No. 1.1;
Matches 17; Conservative 13; Mismatches 27; Indels 17; Gaps 2;

QY 98 NAKKADFKWTW-----GKGMMLAGVVTMAVTRLTEIVLPFTFANNYNRKLKNSLNIALGG 152
:||||||:|
Db 5 SVSSGDYWTWIRQSPGKGLEWIGHY-----YSGNTNPNPSLKSLRTISIDT 52
:||||||:|

QY 153 FEPSFDINMGQASA 166
:||||||:|
Db 53 SKTQFSLKLSVTA 66
:||||||:|

RESULT 10
US-09-134-001C-4175
; Sequence 4175, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4175
; LENGTH: 344
```



```
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4175

Query Match          7.8%; Score 70.5; DB 4; Length 344;
Best Local Similarity 26.6%; Pred. No. 4.9;
Matches 33; Conservative 16; Mismatches 34; Indels 41; Gaps 5;

Qy 4 FLIFLIVFCAFSFAQDDSKSTFNL-----GAGEKFLVYETNKK-----D 44
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 YIIIVLVIV-----YANFANAVENLLNIGSWIVVWGSSWISQIRLQAIKKQGNPD 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 45 SLVPP-----LLNLFGLFGIGSFAQGDILGSLILGLFSDAVGIGLILTGAY 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 264 KLLPYKAPFPVPGVIVLFTLLFLLGSSFGSIASGDILG--VIRNFTPLFLAIYIAH 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 90 LDIK 93
   ||
Db 322 KLIK 325
   ||

RESULT 11
US-09-393-627B-1
; Sequence 1, Application US/09393627B
; Patent No. 6455314
; GENERAL INFORMATION:
; APPLICANT: Wickham, Thomas J.
; APPLICANT: Kovesdi, Imre
; APPLICANT: Roelivink, Petrus W.
; APPLICANT: Bruder, Joseph T.
; TITLE OF INVENTION: Alternatively Targeted Adenovirus
; FILE REFERENCE: 202345
; CURRENT APPLICATION NUMBER: US/09/393,627B
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/099,851
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: US 60/136,529
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.2
; SEQ ID NO 1
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Human adenovirus serotype 5
US-09-393-627B-1

Query Match          7.8%; Score 70.5; DB 4; Length 581;
Best Local Similarity 26.2%; Pred. No. 10;
Matches 48; Conservative 18; Mismatches 72; Indels 45; Gaps 11;

Qy 22 DSKSTFNLGAGEKFLVYETNKKDSLVPFLLNLFGLFGIGSFAQGDILG-----SLI 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 DDLNLTLTATGPGVTINTYSLQTKVG-----ALGFDQGNQNLNVAGGLRIDSONRRLI 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 74 LG-----FDA-----VGIG-LILTGAY-LDIK-----DFDNNAKKADFKWTWGRKM 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 LDVSYFDAQNLRLGQGPLFINSAHNLDINYKGLYFTASNNSKKLEVNLSSTAKGL 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 113 ML-AGVVTMAYTRTEIVLPTFFANNYRNKLNLSNIALG-GFEPSPDINMGQASALGFG 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 330 MFDATAINAGDGLGFGSP-----NAPNTNPLTKTIGHGLE--FDSNKAWVPKLTGTG 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 171 LSF 173
   |||
Db 381 LSF 383
   |||

RESULT 12
US-08-608-241-2
; Sequence 2, Application US/08608241
; Patent No. 5747328
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
```

```
; APPLICANT: Barber, Robert D
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/608,241
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296.93511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-608-241-2

Query Match          7.6%; Score 69; DB 1; Length 376;
Best Local Similarity 25.3%; Pred. No. 8.2;
Matches 37; Conservative 23; Mismatches 54; Indels 32; Gaps 6;

Qy 48 PFLNLFGLFGIG-----SFAQGDILGSLILGLFSDAVGI-----GLILTGAYLIDKDPD 96
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 161 PFDKICYICGVTGTGAVINTAKVEIGAKAVFVGLGGTGLNVIQGLKLAGADMIIGVDL 220
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 97 NNAKKADFKWTWKG-----GMLAGVVTMAYTRTEI-VLPFTFANNYRNK 141
   ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 221 NNAKK-----EMGERFGMTHFVNPSEIDGDVVVAHLVNMTKTDFDQIGGADYTFDCTGNVK 275
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 142 -LKNSLNIALGGFEPSPDINMGQASA 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 276 VMRQALEACHRGWGOSIVIGVAPAGA 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-08-922-182-2
; Sequence 2, Application US/08922182
; Patent No. 5834300
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```


Search completed: March 14, 2003, 14:27:50
Job time : 13.9382 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:25:51 ; Search time 9.94382 Seconds
(without alignments)
820.439 Million cell updates/sec

Title: US-09-508-487-23
Perfect score: 908
Sequence: 1 MNKELIFILVIFCAFSFAQ.....DINMGASALGFLSFKKSY 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues
Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.5	9.7	451	US-09-741-669-422	Sequence 422, Appl
2	87	9.6	413	US-09-910-186A-22	Sequence 22, Appl
3	87	9.6	852	US-10-011-588-25	Sequence 25, Appl
4	81.5	9.0	206	US-09-815-242-11462	Sequence 11462, A
5	78	8.6	544	US-09-738-626-3927	Sequence 3927, Ap
6	78	8.6	544	US-09-948-777-2	Sequence 2, Appli
7	76.5	8.4	417	US-09-738-626-6850	Sequence 6850, Ap
8	74.5	8.2	934	US-08-837-459-19	Sequence 19, Appl
9	73	8.0	406	US-09-738-626-4686	Sequence 4686, Ap
10	73	8.0	421	US-09-738-626-6813	Sequence 6813, Ap
11	72	7.9	670	US-10-060-763-10	Sequence 10, Appl
12	72	7.9	670	US-10-063-763-10	Sequence 10, Appl
13	71.5	7.9	513	US-09-738-626-6860	Sequence 6860, Ap
14	70.5	7.8	95	US-09-187-693-53	Sequence 53, Appl
15	70.5	7.8	410	US-09-767-041-20	Sequence 20, Appl
16	70.5	7.8	579	US-09-870-203A-6	Sequence 6, Appli
17	70.5	7.8	579	US-09-870-203A-8	Sequence 8, Appli
18	70.5	7.8	579	US-09-870-203A-18	Sequence 18, Appl
19	70.5	7.8	579	US-09-870-203A-20	Sequence 20, Appl

20	70.5	7.8	580	10	US-09-870-203A-10	Sequence 10, Appl
21	70.5	7.8	580	10	US-09-870-203A-12	Sequence 12, Appl
22	70.5	7.8	581	10	US-09-758-008-5	Sequence 5, Appli
23	70.5	7.8	581	10	US-09-870-203A-2	Sequence 2, Appli
24	70.5	7.8	581	10	US-09-870-203A-4	Sequence 4, Appli
25	70.5	7.8	581	10	US-09-870-203A-14	Sequence 14, Appl
26	70.5	7.8	581	10	US-09-870-203A-16	Sequence 16, Appl
27	70.5	7.8	600	10	US-09-871-212-5	Sequence 5, Appli
28	70.5	7.8	1161	8	US-08-910-386A-20	Sequence 20, Appl
29	70	7.7	452	10	US-09-815-242-10183	Sequence 10183, A
30	69.5	7.7	650	10	US-09-815-242-13341	Sequence 13341, A
31	68.5	7.5	154	10	US-09-881-752A-118	Sequence 118, App
32	68.5	7.5	475	9	US-09-738-626-6117	Sequence 6117, Ap
33	68	7.5	255	10	US-09-881-752A-106	Sequence 106, App
34	68	7.5	276	9	US-09-895-913A-356	Sequence 356, App
35	68	7.5	570	9	US-09-344-882-28	Sequence 28, Appl
36	68	7.5	570	12	US-10-062-254-367	Sequence 367, App
37	68	7.5	691	10	US-09-841-132-313	Sequence 313, App
38	68	7.5	2039	9	US-10-192-584-7	Sequence 7, Appli
39	68	7.5	2042	9	US-10-192-584-6	Sequence 6, Appli
40	67	7.4	132	10	US-09-800-729-134	Sequence 134, App
41	67	7.4	132	10	US-09-800-729-189	Sequence 189, App
42	67	7.4	156	10	US-09-800-729-135	Sequence 135, App
43	67	7.4	209	9	US-09-813-453A-21	Sequence 21, Appl
44	67	7.4	239	10	US-09-840-787-22	Sequence 22, Appl
45	67	7.4	239	10	US-09-800-729-102	Sequence 102, App

ALIGNMENTS

RESULT 1
US-09-741-669-422
; Sequence 422, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 422
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-422

Query Match	9.7%	Score 88.5;	DB 10;	Length 451;
Best Local Similarity	27.5%	Pred. No. 0.18;		
Matches	30;	Conservative	15;	Mismatches 33;
Indels	31;	Gaps	6;	
QY	46	LVPFLLNLNF---	LGFIGSFAQDILGSLILGFDVAGVIGILTGAYLD-----	IK 93
Db	17	MLPIVILFISKILG-----	MKAGDCKFAGLHIGIFGVIGLVI-GLMLDSIGPAKAAAE	70
QY	94	DFDNNAKADFKW-----	TWCKGMMLAGV-----	VTMAVTRLTEIV 129
Db	71	NFDNLNHHVDVVGCPSPMTWASQIALVAIPAILVNVVAMLLTRTRVV	119	

RESULT 2
US-09-910-186A-22
; Sequence 22, Application US/09910186A
; Publication No. US20030009025A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Research & Material Command

```
; TITLE OF INVENTION: RECOMBINANT VACCINE AGAINST BOTULINUM
; FILE OF INVENTION: NEUROTOXIN
; FILE REFERENCE: A33626-A 067252.0107
; CURRENT APPLICATION NUMBER: US/09/910,186A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: PCT/US00/12890
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: 09/611,419
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/133,865
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,866
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,867
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,868
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,869
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/133,873
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 08/123,975
; PRIOR FILING DATE: 1993-09-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-09-910-186A-22
```

```
Query Match          9.6%; Score 87; DB 9; Length 413;
Best Local Similarity 22.8%; Pred. No. 0.24;
Matches 38; Conservative 31; Mismatches 58; Indels 40; Gaps 7;

QY 37 VYETNKKDS-----LVPEL---LNFLGFGIGSFAQG-DILGGSILILGFD-----AV 79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 162 VIEANKSNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFIAGASILILEFIPPELLIP 221
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 80 GIGLILTGAYLD-----IKDFDNNAKKADFKWTGCKGMLAGVVTMAVTRL----- 125
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 222 VVGAFLESYIDNKNKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQFYTIKEGMVK 281
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 126 -----TEIVLPETFANNYRKLNKLSNIALGGFEPSPFDINMGOA 164
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 282 ALNYQAALIEEIKYRY-NIYSEKEKSNINIDFNDSKLNKLNINQAO 327
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```

```
RESULT 3
US-10-011-588-25
; Sequence 25, Application US/10011588
; Patent No. US20020168727A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Leonard
; TITLE OF INVENTION: RECOMBINANT LIGHT CHAINS OF BOTULINUM
; TITLE OF INVENTION: NEUROTOXINS AND LIGHT CHAIN FUSION PROTEINS FOR USE IN
; TITLE OF INVENTION: RESEARCH AND CLINICAL THERAPY
; FILE REFERENCE: A34796 067252.0113
; CURRENT APPLICATION NUMBER: US/10/011,588
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/910,186
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/611,419
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 60/246,744
; PRIOR FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: 60/311,966
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 25
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant protein encoded by SEQ ID NO:24
US-10-011-588-25
```

```
Query Match          9.6%; Score 87; DB 9; Length 852;
Best Local Similarity 22.8%; Pred. No. 0.56;
Matches 38; Conservative 31; Mismatches 58; Indels 40; Gaps 7;

QY 37 VYETNKKDS-----LVPEL---LNFLGFGIGSFAQG-DILGGSILILGFD-----AV 79
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 601 VIEANKSNTMDKIADISLIVPYIGLALNVGNETAKGNFENAFIAGASILILEFIPPELLIP 660
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 80 GIGLILTGAYLD-----IKDFDNNAKKADFKWTGCKGMLAGVVTMAVTRL----- 125
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 661 VVGAFLESYIDNKNKIITIDNALTKRNEKWSMDYGLIVAQWLSTVNTQFYTIKEGMVK 720
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 126 -----TEIVLPETFANNYRKLNKLSNIALGGFEPSPFDINMGOA 164
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 721 ALNYQAALIEEIKYRY-NIYSEKEKSNINIDFNDSKLNKLNINQAO 766
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```

```
RESULT 4
US-09-815-242-11462
; Sequence 11462, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11462
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11462
```

```
Query Match          9.0%; Score 81.5; DB 10; Length 206;
Best Local Similarity 23.6%; Pred. No. 0.38;
Matches 33; Conservative 21; Mismatches 37; Indels 49; Gaps 6;

QY 60 GSAFGDILGGSILILGFDVAGVIGLITGAYLDIKDFDNNAKKADFKWTGCKGMLAGV-- 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 86 GHFVQGIH-----DAIGV-----TEKIHANQVDFISASETLLILCSEQ 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```



```

QY 121 AVTRLTEIVLPTFANNYNRKLKNSLNTALGGFEP-----S 156
      | | | | | | | | | | | | | | | | | | | |
Db 248 LGTSSSAAITIVT----YQOTLKNDVDVNAVGVVPLCATIHLAGSMKIGLTFEAVFEM 303
      | | | | | | | | | | | | | | | | | | | |
QY 157 FDIINMQOASALGF 169
      | | | | | | | | | | | | | | | | | | | |
Db 304 YDMEVGVGLSIGF 316
      | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-08-837-459-19
; Sequence 19, Application US/08837459
; Patent No. US20020006407A1
; GENERAL INFORMATION:
; APPLICANT: McKee, Marian L.
; APPLICANT: O'Brien, Alison R.
; APPLICANT: Wachtel, Marian R.
; TITLE OF INVENTION: Histidine-Tagged Intimin And Methods Of
; TITLE OF INVENTION: Using Intimin To Stimulate An Immune Response And As An
; TITLE OF INVENTION: Antigen Carrier With Targeting Capability.
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,459
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 04995,0023-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-837-459-19

      Query Match      8.2%  Score 74.5; DB 8; Length 934;
      Best Local Similarity 21.7%  Pred. No. 13;
      Matches 34; Conservative 22; Mismatches 46; Indels 55; Gaps 7;

QY 22 DSKSTFNLGAGEKFLVYETNKKDSLVPFLNLLFLGFGIGSFAQGDILGSLILGFDVAVGI 81
      | | | | | | | | | | | | | | | | | | | |
Db 252 DSREFTANLGAQRE-----FLPANMLGYNV-FIDQDFSGDNLRLG----- 290
      | | | | | | | | | | | | | | | | | | | |

QY 82 GLIITGAYLDIKDFDNNAKKADFKW-TWKGKMLAGVVTMAVTRLTEIVLPFTFANNYNR 140
      | | | | | | | | | | | | | | | | | | | |
Db 291 ---IGGEYW--RDYFKSSVNGYFRMRRWHEYS-----HKKDYDE 324
      | | | | | | | | | | | | | | | | | | | |

QY 141 KLNKSLNALGGFEPSPFDINMQOASALGFLSPFKSY 177
      | | | | | | | | | | | | | | | | | | | |
Db 325 RPANGFDIRFNGLYPSY-----PAUGAKLIYEQYY 354
      | | | | | | | | | | | | | | | | | | | |

RESULT 9

```



```

: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 6813
: LENGTH: 421
: TYPE: prt
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-6813

```

Query Match 8.0%; Score 73; DB 9; Length 421;
Best Local Similarity 25.7%; Pred. No. 6.9;
Matches 36; Conservative 19; Mismatches 47; Indels 38; Gaps 8;

QY 41 NKKDSLVP-----FLNLNLFGLG--GSFAQGDLGGSLILGFDAVGIGLILTGYL 90
| : | | | | | | | | | | | | | | | |
Db 86 NSSGNLIPEETVTLDRFPAGSITGISVGGMGTOGLTGDTADLKSDLLGV----- 136

91 DIKDFDNNAKKADFQWTKWCKGMMLAGVVT---MAVTRLTETIVLPFTFANNYNRKLKNSLN 147
137 -----NEDYMMNMNGVEMNTGRAITQDDVAARPVAVTAPDTF-----NTLFDANPN 182

```
Qy 148 IALGFEPSFDINMGQASAL 167
      :||| | :|:| | | : |
Db 183 LALGS-EVAFELN-GOETFL 200
```

RESULT 11
US-10-060-763-10

```

; Sequence 10, Application US/10060763
; Publication No. US2003002286A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A.J.
; TITLE OF INVENTION: US 2003002286A1el
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/060,763
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0

```

Query Match 7.9%; Score 72; DB 9; Length 670;
Best Local Similarity 21.9%; Pred. No. 15;
Matches 42; Conservative 34; Mismatches 80; Indels 36; Gaps 6;

QY 1 MNKFLIFLVI~~CA~~FSSFAQDDSKSTFNLGAGEKFLVYETNKKDKSLVP~~L~~NLF~~G~~FGIG 60
::: | : | | | | | | | | : | : | | | | | | | | | |

Dd 16 LSKLMPELLAITCAFVS~~K~~TLSGSYNNSMLTOIER-----OFNIPTSLV-----GFING 63

QY 61 SFAQGDILGSLILGFDA-----VGIGLLTGAYLDIKDFDN-NAKKADFKWTWGKGM 112
|| :|:| : | : ||| :| : | : : : |
Db 64 SFEIGNULLIIIFYSGFGTKLRHPIMIGICGVMGLCFLKSLPHFLMNOYEYESTVSVSG 123

QY 113 MLAGVVTMAVTRITEIVLPFTFANNYNRKLNKSL-----NIALGGFEPSFDINMGAS 165
| : : : | : | : : : : | - | : | : ||:
pb 124 NLSSNSFLCMENGTIILPTODPSECTKEVKSI.MVVVVI.VGNIVRG-----MGIIPD 174

QY	166	ALGFGLSFKKSY	177
		: : :	
Db	175	ILPLGISYIEDF	186

RESULT 12
US-10-063-763-10
; Sequence 10, Application US
; Patent No. US20010001663A1

```

; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: No. US20010001663A1el Transporter-Like Genes and Uses Therefo
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/063,763
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: 09/365,162
; PRIOR FILING DATE: 1999-JUL-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10 .
; LENGTH: 670
; TYPE: prt
; ORGANISM: Homo sapiens
; US-10-063-763-10 .

```

Query Match 7.9%; Score 72; DB 12; Length 670;
Best Local Similarity 21.9%; Pred. No. 15;
Matches 42; Conservative 34; Mismatches 80; Indels 36; Gaps 6;

QY 1 MNKFLIFILVIFCAFSFPAQDDSKSTFNLGAGEKFLVYETNKKDSLVPFLNLFLGFGIG 60
::: :|: | || | - | - | - : | - | | |
Dd 16 LSKLKMPELLAITCAFVSKTSLGSYNMSMLTOIER----QFNPTSLV-----GFING 63

QY 61 SFAQGDIILGSLILGFDA-----VGIGLILTAYLDIKDFDN-NAKADFKWTWGKM 112
 || |::| : | : ||| :: | : | : : : |
Db 64 SFEIGNLLLIIFVSYFGTKLHRPIMIGICGVVMGLGFLKSLPHFLMNOYEYESIVSVSG 123

QY 113 MLAGVVTMAVTRLTEIVLPFTFANNYNRKLKNSL-----NIALGGFPSPFDINMQAS 165
| : : | : | : : : : : || - || :
Db 124 NLSSNSFLCMENGTGFIILRPQTDPDSECTKEVKSLMWLVGNVRG-----MGETP 174

Qy	166	ALGFGLSFKSY · 177
		— : : : :
Db	175	ILPLGISYIEDF 186

```

RESULT 13
US-09-7338-626-6860
; Sequence 6860, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6860
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-7338-626-6860

```

Query Match 7.9%; Score 71.5; DB 9; Length 513;
Best Local Similarity 18.6%; Pred. No. 13;
Matches 39; Conservative 28; Mismatches 68; Indels 75; Gaps 6;


```

RESULT 2
TCMA_STRGA          STANDARD;          PRT;          538 AA.
AC P39886;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tetragenomycin C resistance and export protein.
GN TCMA.
OS Streptomyces glaucoscens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinobacteriales; Streptomycetales; Streptomycetaceae; Streptomyces.
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1907;
RN [1]
RP STRAIN=ATCC 13032 / ETH 22794 / GLA.0;
RC MEDLINE=92276347; PubMed=1592819;
TX Guilfoile P.G., Hutchinson C.R.;
TA "Sequence and transcriptional analysis of the Streptomyces glaucoscens tcmAR tetragenomycin C resistance and repressor gene loci."
RT J. Bacteriol. 174:3651-3658(1992).
RL -!- FUNCTION: RESISTANCE TO TETRAGENOMYCIN C BY AN ACTIVE TRANSMEMBRANE C EFFLUX SYSTEM WHICH IS PROBABLY ENERGIZED BY TETRAGENOMYCIN C ELECTROCHEMICAL GRADIENTS.
CC -!- PATHWAY: Polyketide antibiotic tetragenomycin C biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; M80674; AAA67509.1;
PIR; S27687; S27687.
DR InterPro; IPR004638; Efflux_EmrB.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr.1.
DR TIGRFAMs; TIGR00711; efflux_EmrB; 1.
KW Antibiotic resistance; Antibiotic biosynthesis; Transmembrane; Transport.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 239 259 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
FT TRANSMEM 319 339 POTENTIAL.
FT TRANSMEM 342 362 POTENTIAL.
FT TRANSMEM 371 391 POTENTIAL.
FT TRANSMEM 413 433 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
SQ SEQUENCE 538 AA; 54846 MW; DED3F28C1F22AA56 CRC64;
Query Match
Best Local Similarity 9.4%; Score 85; DB 1; Length 538;
Matches 33; Conservative 19; Mismatches 38; Indels 44; Gaps 6;
OY 49 PFGLNLFLGFGIG-----SFAQGDILGGFLI--LGFDA-----VCGIGLILTGAVLDIKA 95
DB 145 PKLNLWAIQISWGVCAGTAAGPIIGLLVQHVGHVEAFINVPVGLAALVAGLVILTD 204
-----
OY 96 LDKNAPKA-----AFKWTWKG-----GMMLAGAV-----TMAYT 124
DB 205 RAERAPKSFDSGIVLLSGAMFLVWGLIKAPAWGWDLRTLGLFLAAVLAFAFTLRES 264
OY 125 RLTEIIIPFTFANS 138
DB 265 RATEPLMLAMFRS 278
RESULT 3
DTA2_RHILO
ID DTA2_RHILO STANDARD; PRT; 442 AA.
AC Q986R8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C4-dicarboxylate transport protein 2.
GN DCTA2 OR MLL7237.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S., Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpou S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
CC -!- FUNCTION: Responsible for the transport of dicarboxylates such as succinate, fumarate, and malate from the periplasm across the membrane. This transport system plays an important role in the energy supply of rhizobium-legume symbionts (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY (SDF, TC 2.A.23).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; AP003011; BAB53385.1;
DR InterPro; IPR001991; Na/dico_symp.
DR Pfam; PF00375; SDF; 1.
DR PRINTS; PR00173; EDTNSPORT.
DR PROSITE; PS00713; NA_DICARBOXYL_SYMP_1; 1.
DR PROSITE; PS00714; NA_DICARBOXYL_SYMP_2; 1.
KW Transport; Sugar transport; Transmembrane; Inner membrane; Symport; Complete proteome.
FT TRANSMEM 20 39 POTENTIAL.
FT TRANSMEM 52 74 POTENTIAL.
FT TRANSMEM 89 111 POTENTIAL.
FT TRANSMEM 141 158 POTENTIAL.
FT TRANSMEM 162 179 POTENTIAL.
FT TRANSMEM 200 221 POTENTIAL.
FT TRANSMEM 231 253 POTENTIAL.
FT TRANSMEM 342 364 POTENTIAL.
FT TRANSMEM 368 387 POTENTIAL.
SQ SEQUENCE 442 AA; 46094 MW; 75704505629BD82 CRC64;
Query Match
Best Local Similarity 9.2%; Score 83.5; DB 1; Length 442;
Matches 51; Conservative 24; Mismatches 86; Indels 47; Gaps 10;

```

```

RESULT 5
PHSC_AQUAE STANDARD; PRT; 692 AA.
ID PHSG_AQUAE STANDARD; PRT; 692 AA.
AC 066932;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycogen phosphorylase (EC 2.4.1.1).
GN GLGP OR AQ_717.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5.
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."
RL Nature 392:353-358(1998).
CC -1- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN
CC CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
CC THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
CC PROPERTIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: [(1,4)-alpha-D-glucosyl](N) + phosphate =
CC [(1,4)-alpha-D-glucosyl](N-1) + alpha-D-glucose 1-phosphate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000704; AAC06896.1; -
DR HSSP; P00489; 3AMV.
DR InterPro; IPR000811; GT_35.
DR Pfam; PF00343; phosphorylase; 1.
DR PROSITE; PS0102; PHOSPHORYLASE; 1.
DR Transfaser; Glycosyltransferase; Carbohydrate metabolism;
KW Glycogen metabolism; Pyridoxal phosphate; Complete proteome.
FT BINDING 586 586 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 692 AA; 81158 MW; 4DC0EA2E2E58BEA4 CRC64;
-----
Query Match 9.0%; Score 81.5; DB 1; Length 692;
Best Local Similarity 24.6%; Pred No. 7.1;
Matches 44; Conservative 17; Mismatches 49; Indels 69; Gaps 10;

QY 39 YETSKDPVYFPL-----LNLFLGFGISFAQGDILGLGFLILGFDVAVGIGLILTGA 89
|| : ||| || : : : ||| ||| ||| : : :
DB 84 YEDTYKRPV-FLSPYGLHHTLIYAG-GIG-FLAGDILKESDGLGFLIGVGMYPQG 140
|| : ||| || : : : ||| ||| ||| : : :
QY 90 Y-----LDI-----KALDKNAPKAFK-----TWGK 111
| ||| ||| |||
DB 141 YVKORIRVDGQWEDQAQNKELMPVKVLVDKEG-----KWLKCYVYVRDEKVFYGV-- 193
|
QY 112 GMLLAGAVTAVTRLTETIIIPFTFANSYNRKLNKSLNIAFGGFPFSDINQASALGF 170
|| : ||| : : : ||| : : : ||| : : :
DB 194 -----EVNKGKTKLYLLDTNVEENTPWNREISRL-----YVPDKDLRLRQIVLGF 240
|

RESULT 6
BXB_CLOBO STANDARD; PRT; 1290 AA.
DB BXB_CLOBO STANDARD; PRT; 1290 AA.

```

AC P10844; P10843;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BoNT/B)
 DE (Bontoxilysin B).
 OS BOTB.
 GN Clostridium botulinum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 OX (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92384550; PubMed=1514783;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
 RA Minton N.P.;
 RT "Molecular cloning of the Clostridium botulinum structural gene
 RT encoding the type B neurotoxin and determination of its entire
 RT nucleotide sequence.";
 RT Appl. Environ. Microbiol. 58:2345-2354(1992).
 RN (2)
 RP SEQUENCE OF 35-245 FROM N.A.
 RC STRAIN=NCTC 7273;
 RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE OF 633-993 FROM N.A.
 RC STRAIN=NCTC 7273;
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulin neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 RN (4)
 RP SEQUENCE OF 1-44 AND 441-466.
 RC STRAIN=657;
 RX MEDLINE=69000987; PubMed=3139097;
 RA Dasgupta B.R., Datta A.;
 RT "Botulinum neurotoxin type B (strain 657): partial sequence and
 RT similarity with tetanus toxin.";
 RL Biochimie 70:811-817(1988).
 RN (5)
 RP SEQUENCE OF 1-16 AND 441-458.
 RC STRAIN=OKRA;
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Sathyanarthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN (6)
 RP IDENTIFICATION AS ZINC-PROTEASE.
 RX MEDLINE=93054694; PubMed=1429690;
 RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
 RT "Botulinum neurotoxins are zinc proteins.";
 RL J. Biol. Chem. 267:23479-23483(1992).
 RN (7)
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=93063293; PubMed=1331807;
 RA Schiavo G., Benfenati F., Poulin B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin.";
 RL Nature 359:832-835(1992).
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-|-PHE-77 BOND OF
 CC SYNAPTOBREVIN-2.
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.

CC -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIANTE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81186; AAA23211.1; -;
 CC EMBL; Z11934; CAA77991.1; -;
 CC EMBL; X70817; CAA50148.1; -;
 CC PIR; S07128; S07128.
 CC PIR; S07155; S07155.
 CC PIR; S08562; S08562.
 CC PIR; S08573; S08573.
 CC PIR; S08574; S08574.
 CC PIR; A48940; A48940.
 CC HSSP; P10845; 3BTA.
 CC MEROPS: M27.002; -;
 CC InterPro: IPR000395; Bontoxilysin.
 CC InterPro: IPR000130; Zn_MTPeptidse.
 CC Pfam: PF01742; Peptidase_M27; 1.
 CC PRINTS; PR00760; BONTOXILYSIN.
 CC ProDom: PD001963; Bontoxilysin; 1.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 KW INIT_MET 0
 FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
 FT METAL 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
 FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 436 445 INTERCHAIN (PROBABLE).
 FT CONFLICT 29 29 T -> M (IN REF. 4).
 FT CONFLICT 217 217 R -> G (IN REF. 2).
 FT CONFLICT 224 224 A -> S (IN REF. 2).
 FT CONFLICT 463 463 S -> R (IN REF. 4).
 SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;
 Query Match 9.0%; Score 81.5; DB 1; Length 1290;
 Best Local Similarity 23.8%; Pred. No. 13;
 Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;
 QY 47 IVPFLNLFGLFGIGSFAOGDILGFLILGFDV-----GIGLILTGAYLD--- 92
 Db 619 IVPY-IGLALNVG-NETAKGNFENAFIAGASILLFIPPELLIPVVGAFLESYIDNKK 676
 QY 93 -IKALDNAPKAAFKWTKGKMLAGAVTMAVTRL-----TEIIPFT 134
 Db 677 IKTIDNALTKRNEKWSMDYGLIVQWLSVTNTQFYTIKRGMYKALNYQAQALEIIKYR 736
 QY 135 FANSYWRKLNKSLNIAFGGFEPSFDINMQQA 165
 Db 737 Y-NIYSEKESKNIDFNDINSKLNENGINQA 766
 RESULT 7
 YQGP_BACSU
 ID YQGP_BACSU STANDARD; PRT; 507 AA.
 AC P54493;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yqgp.

```

FT TRANSMEM 312 332 POTENTIAL.
FT TRANSMEM 365 POTENTIAL.
SQ SEQUENCE 507 AA; 56439 MW; 7FDF65D4EB2C81E5 CRC64;

Query Match          9.0%; Score 81; DB 1; Length 507;
Best Local Similarity 28.3%; Pred. No. 5.8; Indels 22; Gaps 5;
Matches 30; Conservative 17; Mismatches 37;

QY   3 KFLIVL-----LAFCVESFAQADDSAFN-LCA-----GEKLLAYETSKKDPIV 48
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |
Db 260 RFLLIYLAGITGSIASFVPSPSAGASGAIFCGGLGALLYVALSNRKMFLRTIGTNIIV 319
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |

QY   49 PFLNLFLFGIGSFQAQ-----GDILGGLFLILGFDAVGIGLILTGAY 90
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |
Db 320 IIIINLGFGFAVENIDNSGHIGGLIGCF----FAAALGPLKAGAF 361
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |

RESULT 8
OOXB_AGRt4 STANDARD; PRT; 371 AA.
ID OOXB_AGRt4
AC Q59159;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Opine oxidase subunit B (EC 1.-.-.) (Octopine oxidase subunit B).
GN OOXB.
OS Agrobacterium tumefaciens (strain Ach5).
OG Plasmid pTiAch5
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
NCBI_TaxID=176298;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94321320; PubMed=8045881;
RA Zanker H., Lurz G., Langridge U., Langridge P., Kreusch D.,
RA Schroeder J.;
RA "Octopine and nopaline oxidases from Ti plasmids of Agrobacterium
RT tumefaciens: molecular analysis, relationship, and functional
RT characterization.";
RL J. Bacteriol. 176:4511-4517(1994).
CC -!- FUNCTION: OXIDATIVE CLEAVAGE OF OCTOPINE INTO L-ARGININE AND
CC PYRUVATE (BY SIMILARITY).
CC -!- PATHWAY: CATABOLIC UTILIZATION OF OCTOPINE.
CC -!- SUBUNIT: HETERODIMER OF A SUBUNIT AND A B SUBUNIT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licensese@isb-sib.ch)
CC -----
EMBL; Z30328; CAAB2987.1; .
DR InterPro; IPR000927; D_aa_oxidase.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF01266; DAO; 1.
DR Oxidoreductase; Plasmid.
KW OXIDOREDUCTASE; PLASMID.
SQ SEQUENCE 371 AA; 39939 MW; 948705136822AC52 CRC64;

Query Match          8.9%; Score 80.5; DB 1; Length 371;
Best Local Similarity 30.9%; Pred. No. 4.7;
Matches 38; Conservative 13; Mismatches 39; Indels 33; Gaps 6;

QY   68 ILGGFLILGFDAVGIG-----LILTQAYLDIKALDKNAKPAKFWTWKGMMML----- 115
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |
Db 8 IIIGGLYGASTAWGLARSQTKPLVDGADLDLRASN---FALVVYQGKLGHAPHYALW 64
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |

QY   116 --AGA---VTMAYTRLTEII-----PTFFANS-----YNRKLSNLIAGFGPEPS 157
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |
Db 65 SDASAREWRPMTANFLDDSGIDVGLQQDGAPTFALSBEELFANQDMESIELETNGRAPQ 124
      :| | : | :| | : | :| | : | :| | : | :| | : | :| |
OV 158 FDI 160

```


CCDA_BACSU	STANDARD;	PRT;	235 AA.
AC	CCDA_BACSU	STANDARD;	PRT;
ID	P45706;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Cytochrome c-type biogenesis protein ccda.		
DE	CCDA.		
GN			
OC	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.		
OX	NCBI_TaxID=1423;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=168;		
RC	MEDLINE=97124194; PubMed=8969507;		
RX	Rose M., Entian K.D.;		
RT	"New genes in the 170 degrees region of the Bacillus subtilis genome		
RT	encode DNA gyrase subunits, a thioredoxin, a xylanase and an amino		
RT	acid transporter";		
RL	Microbiology 142:3097-3101(1996).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RN	STRAIN=168;		
RC	MEDLINE=97221596; PubMed=9068642;		
RX	Schioett T., von Wachenfeldt C., Hederstedt L.;		
RA	"Identification and characterization of the ccda gene, required for		
RT	cytochrome c synthesis in Bacillus subtilis.";		
RT	J. Bacteriol. 179:1962-1973(1997).		
RL	[3]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=168;		
RC	MEDLINE=98044033; PubMed=9384377;		
RX	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,		
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,		
RA	Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,		
RA	Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,		
RA	Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,		
RA	Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,		
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,		
RA	Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,		
RA	Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,		
RA	Gulseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,		
RA	Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,		
RA	Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,		
RA	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,		
RA	Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,		
RA	Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,		
RA	Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,		
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudesa B., Park S.H.,		
RA	Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,		
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,		
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,		
RA	Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,		
RA	Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,		
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,		
RA	Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,		
RA	Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,		
RA	Viari A., Wambutt R., Wedler H., Wedler H., Weitzenecker T.,		
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,		
RA	Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;		
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus		
RT	subtilis.";		
RT	Nature 390:249-256(1997).		
RL	[4]		
RN	FUNCTION, AND SUBCELLULAR LOCATION.		
RP	STRAIN=168;		
RX	MEDLINE=20245543; PubMed=10781554;		
RA	Schioett T., Hederstedt L.;		
RT	"Efficient spore synthesis in Bacillus subtilis depends on the Ccda		
RT	protein.";		
RL	J. Bacteriol. 182:2845-2854(2000).		
CC	-1- FUNCTION: REQUIRED FOR CYTOCHROME C SYNTHESIS AND STAGE V OF		
CC	SPOULATION. MIGHT TRANSFER REDUCING EQUIVALENTS ACROSS THE		

```
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY
CC (SDF, TC 2.A.23).
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z11529; CAA77618.1; .
CC PIR; S25701; S25701.
CC PIR; S27384; S27384.
CC InterPro; IPR001991; Na/dico_symp.
CC Pfam; PF00375; SDF; 1.
CC PROSITE; PS00713; NA_DICARBOXYL_SYM_1; 1.
CC PROSITE; PS00714; NA_DICARBOXYL_SYM_2; 1.
CC Transport; Sugar transport; Transmembrane; Inner membrane; Symport.
CC TRANSMEM 18 40 POTENTIAL.
CC FT TRANSMEM 53 75 POTENTIAL.
CC FT TRANSMEM 90 112 POTENTIAL.
CC FT TRANSMEM 142 159 POTENTIAL.
CC FT TRANSMEM 163 180 POTENTIAL.
CC FT TRANSMEM 201 222 POTENTIAL.
CC FT TRANSMEM 232 254 POTENTIAL.
CC FT TRANSMEM 327 349 POTENTIAL.
CC FT TRANSMEM 364 386 POTENTIAL.
CC SEQUENCE 444 AA; 46044 MW; AA2A5EA9B6780653 CRC64;

Query Match
Best Local Similarity 29.1%; Score 77.5; DB 1; Length 444;
Matches 44; Conservative 15; Mismatches 47; Indels 45; Gaps 9;

QY 38 AYETSKDPVIFPLNLFGLFGIGSFAQGDILGGFLGDAV-GIGLILTG----- 88
Db 137 AHEQS-----IVGFLNIITPTTVGAFADGDIL---QVLFFSVLFGIALAMVGEKGEQVYN 189
QY 89 -----AYLDIKALDKNAKRAAF---KWTWCK-GM-----MLAGAVTNVAVTLRLTEI 129
Db 190 FLNLTAPVFKLVAILMKRAAPGAFGAMAFITGKYGVSIGANLMLIG--TFYITSLLEFV 247
QY 130 IIPFTFANSYN-----RLKNSLNIAFG 152
Db 248 FIVIGAVARYNGFSIVALLRYIKEELLVIG 278

RESULT 15
SLAP_CAMPE
ID SLAP_CAMPE STANDARD; PRT; 933 AA.
AC P35827;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE S-layer protein (Surface array protein) (SAP).
GN SAP.
OS Campylobacter fetus.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=196;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC STRAIN=84-32 / 23D;
RX MEDLINE=90354448; PubMed=2387868;
RA Blaser M.J., Gotschlich E.C.;
RT "Surface array protein of Campylobacter fetus. Cloning and gene
RT structure."
RL J. Biol. Chem. 265:14529-14535(1990).
RN [2]
RP ERRATUM.
RX MEDLINE=91035477; PubMed=2229082;
```

```
RA Blaser M.J., Gotschlich E.C.;
RL J. Biol. Chem. 265:19372-19372(1990).
CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS PROTEIN IS
CC CRITICAL FOR VIRULENCE.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J05577; AAA23032.1; .
CC Cell wall; S-layer.
CC SEQUENCE 933 AA; 96757 MW; F88C729B4BA5B1E9 CRC64;

Query Match
Best Local Similarity 8.5%; Score 76.5; DB 1; Length 933;
Matches 42; Conservative 21; Mismatches 73; Indels 35; Gaps 6;

QY 14 VFSSFAQADDSKSAFN---LGAGEKLLAYETSKDPIVPFLNLFGLFGIGSFAQGDIL 69
Db 615 ISDAVATADLSSSAFNKNSVIITTKAAADTTLTKDQVIN-----FTAADAGSVK----- 664
QY 70 GGFLILGFDVAVGIGLILTGAYLDIKALDKN-----APKAAFKWTWGMKMLAGAVTMA 122
Db 665 --LITVKLNDVTALMIVVKIVLDAAAKDTNIALGTAADKAL-----VIDTGTETLN 714
QY 123 VTRLTEIILPFTFANSYNKLNLSNIAFGGPEPSFDINWGASALGFELS 173
Db 715 IISLVKATSPETTANIVNAKLDTVTSIIIDGMQ-----ITLGRAGTAGTDYS 761
```

Search completed: March 14, 2003, 14:24:56
Job time : 10.3333 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:22:51 ; Search time 24.3333 Seconds
(without alignments)
1507.249 Million cell updates/sec

Title: US-09-508-487-21

Perfect score: 904

Sequence: 1 MNKFLIVLLAFVCFSSFAQ.....DINMGQASALGFELSPFKKSY 178

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	904	100.0	178	2 Q9X3U9	Q9x3u9 borrelia af
2	786.5	87.0	177	2 Q9X3V0	Q9x3v0 borrelia ga
3	769.5	85.1	179	16 O51065	O51065 borrelia bu
4	293.5	32.5	161	16 O50896	O50896 borrelia bu
5	285	31.5	170	16 O50885	O50885 borrelia bu
6	240	26.5	190	16 O50696	O50696 borrelia bu
7	237	26.2	161	2 Q9S011	Q9s011 borrelia bu
8	160	17.7	95	2 O31325	O31325 borrelia bu
9	95.5	10.6	574	16 O51143	O51143 borrelia bu
10	92	10.2	427	2 Q93RE6	Q93re6 vibrio para
11	90	10.0	325	16 Q8REW2	Q8rew2 fusbacteri
12	90	10.0	479	2 Q9FDM0	Q9fdm0 zymomonas m
13	89.5	9.9	469	16 Q97JE7	Q97je7 clostridium
14	88	9.7	312	16 Q9RTD0	Q9rtd0 deinococcus
15	87.5	9.7	616	16 Q8UAV6	Q8uav6 agrobacteri
16	85.5	9.5	482	16 Q9ZMT7	Q9zmt7 helicobacte

17	85.5	9.5	538	17	Q9HHI2	Q9hh12 halobacteri
18	85	9.4	363	16	Q929N4	Q929n4 listeria in
19	85	9.4	364	16	Q8Y5D0	Q8y5d0 listeria mo
20	84	9.3	1606	3	Q9UW87	Q9uw87 candida alb
21	83.5	9.2	650	16	Q97RE5	Q97re5 streptococc
22	83.5	9.2	709	16	Q73100	Q73100 synecocyst
23	83	9.2	445	16	Q9KBP3	Q9kbp3 bacillus ha
24	82.5	9.1	559	5	Q24796	Q24796 echinococcu
25	82.5	9.1	559	5	Q05768	Q05768 echinococcu
26	81.5	9.0	442	17	Q976V7	Q976v7 sulfolobus
27	81	9.0	214	16	O55742	O55742 synecocyst
28	81	9.0	305	16	O25355	O25355 helicobacte
29	80.5	8.9	357	16	O66804	O66804 aquifex ao
30	80.5	8.9	661	10	Q9XGQ5	Q9xgq5 oryza sativ
31	80	8.8	183	16	Q9XIG6	Q9xig6 thermotoga
32	80	8.8	336	16	P73160	P73160 synecocyst
33	80	8.8	379	16	Q92L18	Q92l18 rhizobium m
34	80	8.8	518	16	Q8XZV3	Q8xzv3 ralstonia s
35	80	8.8	779	16	Q92FB1	Q92fb1 listeria in
36	79.5	8.8	419	2	Q933G9	Q933g9 pseudomonas
37	79.5	8.8	517	16	Q8XZC8	Q8xzc8 ralstonia s
38	79.5	8.8	559	5	O24788	O24788 echinococcu
39	79.5	8.8	904	12	Q9IWA2	Q9iwa2 wheat yello
40	79	8.7	298	2	Q9FB74	Q9fb74 helicobacte
41	79	8.7	399	2	Q9ANY6	Q9anv6 lactobacill
42	79	8.7	419	2	Q9KJY5	Q9kly5 klebsiella
43	79	8.7	429	16	Q9KFP6	Q9kfp6 bacillus ha
44	79	8.7	433	17	Q8U2X3	Q8u2x3 pyrococcus
45	79	8.7	539	10	O80918	O80918 arabisopsi

ALIGNMENTS

RESULT 1

Q9X3U9 ID Q9X3U9 PRELIMINARY: PRT; 178 AA.
AC Q9X3U9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Membrane protein P13.
GN P13.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_taxid=29518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACAI;
RX MEDLINE=21189251; PubMed=11292755;
RA Noppa L., Ostberg Y., Lavrinovich M., Bergstrom S.;
RT "P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains."
RL Infect. Immun. 69:3323-3334(2001).
DR EMBL: AF085740; AAD28361.1; -;
SQ SEQUENCE 178 AA; 19194 MW; C1C032EB55320C86 CRC64;

Query Match 100.0%; Score 904; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 5.2e-74;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPVPPFLNLFGLFGCI 60
|||||

Db 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNLGAGEKLLAYETSKDPVPPFLNLFGLFGCI 60
|||||

Qy 61 GSFAQGDILGGFLILGFDVAVGIGLITGAYLIDIKALDKNAPKAFKWTGKGMMLAGAVT 120
|||||

Db 61 GSFAQGDILGGFLILGFDVAVGIGLITGAYLIDIKALDKNAPKAFKWTGKGMMLAGAVT 120
|||||

Qy 121 MAVTFLTEIIIPFTFANSYNRKLNLSIAFGGFPSPDINNMQASALGFELSPFKKSY 178
|||||

Db 121 MAVTFLTEIIIPFTFANSYNRKLNLSIAFGGFPSPDINNMQASALGFELSPFKKSY 178
|||||

```
RESULT 2
Q9X3V0
ID Q9X3V0 PRELIMINARY; PRT; 177 AA.
AC Q9X3V0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Membrane protein P13.
GN P13.
OS Borrelia garinii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP90;
RX MEDLINE=21189251; PubMed=11292755;
RA Noppa L., Ostberg Y., Lavrinovitcha M., Bergstrom S.;
"p13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains.";
RL Infect. Immun. 69:3323-3334(2001).
RL EMBL: AF085741; AAD28362.1; -.
SQ SEQUENCE 177 AA; 19308 MW; 429928CC426E1111 CRC64;

Query Match 87.0%; Score 786.5; DB 2; Length 177;
Best Local Similarity 86.5%; Pred. No. 2.1e-63;
Matches 154; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNILGAGEKLLAYETSKKDPVFPFLNLFGLGFI 60
DB 1 MNKFLIFVLCFVCFSSFAQADDSKSTFNIGAGEKELVYETNKKDSLVPPFLNLFGLGFI 59
QY 61 GSFAQGDILGGFLILGFDVAGVIGLITGAYLDIKALDKNAPKAAFWTKWGKGMMLAGAVT 120
DB 60 GSFAQGDILGGSLILGFDVAGVIGLITGAYLDIKDFDNNAKADKFWTKWGKGMMLAGVVT 119
QY 121 MAVTRLTEIIPFTFANSYNRKLNLSNIAFGFEPFSDINMGQASALGFELSPKSY 178
DB 120 MAVTRLTEIVLPFTFANNYNRKLNLSNIALGGFEPFSDINMGQASALGFLSPKSY 177

RESULT 3
O51065
ID O51065 PRELIMINARY; PRT; 179 AA.
AC O51065;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BB0034.
GN BB0034 OR P13.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
"Genomic sequence of a Lyme disease spirochete."
RL Nature 390:580-586(1997).
RL EMBL: AE000790; AAC66226.1; -.
DR TIGR: BBA01;
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 161 AA; 17887 MW; D93BDB326FE2DA30 CRC64;

Query Match 32.5%; Score 293.5; DB 16; Length 161;
Best Local Similarity 41.1%; Pred. No. 6.1e-19;
Matches 74; Conservative 24; Mismatches 61; Indels 21; Gaps 6;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNILGAGEKLLAYETSKKDPVFPFLNLFGLGFI 60
DB 1 MKKIETLILIFGLTIQIFA-TKDTQNRKIEGI-ESFNKYDKRKNPVGPELNLFLPFGI 58
QY 61 GSFAQGDILGGFLILGFDVAGVIGLITGAYLDIKALDKNAPKAAFWTKWGKGMMLAG-AV 119
DB 59 GSFAQGDILGGSVLGGFLNLGAILGTGTLN-----HRETOLTGTILGVGA 106
QY 120 TMAVTRLT-EIIPFTFANSYNRKLNLSNIAFGFEPFSDINMGQASALGFELSPKSY 178
DB 107 SWLTSTYSLTIIPFTFANHNNENLRKRSALAGFEPFDLGIN-----GFOLSPKSY 161
```

```
Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AE001117; AAC66426.1; -.
DR EMBL: AF085739; AAD28360.1; -.
DR TIGR: BB0034; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 179 AA; 19104 MW; FD6056E8E24E5D5D CRC64;

Query Match 85.1%; Score 769.5; DB 16; Length 179;
Best Local Similarity 84.9%; Pred. No. 7.3e-62;
Matches 152; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNKFLIVLLAFVCFSSFAQADDSKSAFNILGAGEKLLAYETSKKDPVFPFLNLFGLGFI 59
DB 1 MNKLLIFVLATFCVCFSSFAQADDSKNGAFGMSAGEKLVVETSKQDDPVPFLLNLFGLGFI 60
QY 60 GSFAQGDILGGFLILGFDVAGVIGLITGAYLDIKALDKNAPKAAFWTKWGKGMMLAGAV 119
DB 61 GSFAQGDILGGSLILGFDVAGVIGLITGAYLDIKALDKGTTKKAFFQMTWKGVMYLAGVV 120
QY 120 TMAVTRLTEIIPFTFANSYNRKLNLSNIAFGFEPFSDINMGQASALGFELSPKSY 178
DB 121 TMAVTRLTEIIPFTFANSYNRKLNLSNIALGGFEPFSDVAMGQSSALGFELSPKSY 179

RESULT 4
O50896
ID O50896 PRELIMINARY; PRT; 161 AA.
AC O50896;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBA01.
GN BBA01.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp54
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
Smith H.O., Venter J.C.;
"Genomic sequence of a Lyme disease spirochete."
RL Nature 390:580-586(1997).
RL EMBL: AE000790; AAC66226.1; -.
DR TIGR: BBA01;
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 161 AA; 17887 MW; D93BDB326FE2DA30 CRC64;
```

```
RESULT 5
O50885 ID O50885 PRELIMINARY; PRT; 170 AA.
AC O50885;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBI31.
GN BBI31.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Plasmid lp28-4.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi."
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 170 AA; 18227 MW; 0E71B633EA80FB22 CRC64;

Query Match 31.5%; Score 285; DB 16; Length 170;
Best Local Similarity 38.8%; Pred. No. 3.8e-18;
Matches 69; Conservative 25; Mismatches 70; Indels 14; Gaps 3;

QY 1 MNKFLIVLLAFCVFSSFAQADSKSAFNLGAGEKLLAYETSKDPVLPFLNLFGLGFI 60
DB 7 MKKLTILFLSLTIQIATQD--KLEKSVGSIIETIMKYSEKATILAPFLNLFGLTGI 64

QY 61 GSFAQGLDGLFGLFDVAGIGILTCAYLDIKALOKNAPKAAFKWTGKGMMLAGAVT 120
DB 65 GSFVQGDYIGGAVLGSQGLGILCIAGNLTGHTDDETRA-----TTGHIITIGVT 117

QY 121 MAVTRLEIIPFTFANSYNRKLNSLNIAFGPEPSPFDINMGASALGFELSPKSY 178
DB 118 IIAASHIASLIIPFTFANKNANLKRLLGIDTAGEPNFDIGIS-----GFQLSPKRY 170

RESULT 6
O50696 ID O50696 PRELIMINARY; PRT; 190 AA.
AC O50696;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BBH41.
GN BBH41.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Plasmid lp28-3.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi."
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 170 AA; 18227 MW; 0E71B633EA80FB22 CRC64;

Query Match 31.5%; Score 285; DB 16; Length 170;
Best Local Similarity 38.8%; Pred. No. 3.8e-18;
Matches 69; Conservative 25; Mismatches 70; Indels 14; Gaps 3;

QY 1 MNKFLIVLLAFCVFSSFAQADSKSAFNLGAGEKLLAYETSKDPVLPFLNLFGLGFI 60
DB 7 MKKLTILFLSLTIQIATQD--KLEKSVGSIIETIMKYSEKATILAPFLNLFGLTGI 64

QY 61 GSFAQGLDGLFGLFDVAGIGILTCAYLDIKALOKNAPKAAFKWTGKGMMLAGAVT 120
DB 65 GSFVQGDYIGGAVLGSQGLGILCIAGNLTGHTDDETRA-----TTGHIITIGVT 117

QY 121 MAVTRLEIIPFTFANSYNRKLNSLNIAFGPEPSPFDINMGASALGFELSPKSY 178
DB 118 IIAASHIASLIIPFTFANKNANLKRLLGIDTAGEPNFDIGIS-----GFQLSPKRY 170

RESULT 7
O9S011 ID O9S011 PRELIMINARY; PRT; 161 AA.
AC O9S011;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Conserved hypothetical protein.
GN BB006.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Plasmid lp56.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Lathigra R.,
RA White O., Dodson R., Hickey E.K., Gwinn M., Peterson J., van Vugt R.,
RA Palmer N., Haft D., Rosa P., Stevenson B.;
RT "A bacterial genome in flux: The twelve linear and nine circular
extrachromosomal DNAs in an infectious isolate of the Lyme disease
spirochete Borrelia burgdorferi."
RL Mol. Microbiol. 0:0-0(1999).
DR EMBL; AE001584; AAF07707.1;
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 161 AA; 17016 MW; 14D4906CC8107CD2 CRC64;

Query Match 26.2%; Score 237; DB 2; Length 161;
Best Local Similarity 39.7%; Pred. No. 7.7e-14;
Matches 60; Conservative 24; Mismatches 49; Indels 18; Gaps 6;

QY 18 FAQADDSKSAFNLGAGE--KLLAYETSKDPVLPFLNLFGLGIGSFAQGDILGFFIL 75
DB 4 FAQEKLEK-----GVGDIATVMKYSEKATILAPLLNLFSLGIGSFVQGDYIGGALL 58

QY 76 GFDAVGIGLITGAYL--DIKALOKNAPKAAFKWTGKGMMLA--GATMAVTRLEI 131
DB 59 GSQVLG--GILIMAGYMTGDIGFVTESTA-----TVITGGVLSGIGGLTAAIYITGIII 111

QY 132 PFTFANSYNRKLNSLNIAFGPEPSPFDINM 162
DB 112 PFKFANRYNADLKKRLGIALAGLEPNFDIGI 142

RESULT 8
O31325
```

```

ID O31325 PRELIMINARY; PRT; 95 AA.
AC O31325;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE (clone 8) s3.
GN S3.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N40;
RX MEDLINE=96305348; PubMed=8688460;
RA Feng S., Das S., Barthold S.W., Fikrig E.;
RT "Characterization of two genes, p11 and p5, on the Borrelia
  burgdorferi 49-kilo base linear plasmid."
  Biochim. Biophys. Acta 1307:270-272(1996).
  EMBL; L41151; AAB62281.1; -.
  SEQUENCE 95 AA; 10739 MW; D12F6791F8920C3A CRC64;

Query Match 17.7%; Score 160; DB 2; Length 95;
Best Local Similarity 44.4%; Pred. No. 3.8e-07;
Matches 36; Conservative 12; Mismatches 31; Indels 2; Gaps 2;

QY 1 MNKFLIVLLARFCVSSFAQADDSKSAFNLAGKLLAYETSKDPIVFPFLNLFGLFGI 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  1 MKKIFLILIFGLTIEIFA-TRDTQNRIEKG-ESFNKYDKEKNPIGPFLLNLFPLFGI 58
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 GSFAGDILGGFLIIGFDVANG 81
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  59 GSFVQGDYIGGGVILGFLNLLG 79

RESULT 9
OS1143 PRELIMINARY; PRT; 574 AA.
AC OS1143;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PTS system, maltose and glucose-specific IIAB component (MALX).
GN BB0116.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Castlens S., Huang W.M., Sutton G.G., Clayton R.A.,
  Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
  Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
  Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
  van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
  Utterback T., Wattley L., McDonald L., Artiach P., Bowman C.,
  Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
  Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
  burgdorferi."
  Nature 390:580-586(1997).
  EMBL; AE001124; AAC66516.1; -.
  DR HSP; P05053; 1IBA.
  DR TIGR; BB0116; -.
  DR InterPro; IPR001996; PTS_EIIB.
  DR InterPro; IPR003352; PTS_EIIC.
  DR Pfam; PF00367; PTS_EIIB; 1.
  DR Pfam; PF02378; PTS_EIIC; 1.
  KW Complete proteome.
  SQ SEQUENCE 574 AA; 63513 MW; 5CB04166E9DA8D67 CRC64;

Query Match 10.6%; Score 95.5; DB 16; Length 574;
Best Local Similarity 25.5%; Pred. No. 1.9;

```

```

Matches 55; Conservative 25; Mismatches 67; Indels 69; Gaps 12;

QY 6 IVVLLAFCVF-----SSFAQADDSKSAFNLAG-----GKLLAYETSKDPI--- 47
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  188 IVIILPFCVFLAIFCLIWSSF---DDLIALGLFVRFEYFGSFLYGLNRLLLPLGLH 244
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  48 ----VPF-----LLNLF-----LGFIGSFAQG--DILGGF-LILGFD 78
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  245 SILSPFPEFTSLGGVEIVNGDVRLGNIFYAQLDPSLKGKSSGFAKISSGFYLSIMEG 304
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  79 AVGIGLILTCAYLDIKALDKNAKPAKWTWKGMMLAGAVTMVTRLEIILPPTFANS 138
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  305 LFGAAL---GVYKGVIVHEDKNVAA-----LLFSGALTAFLTGITE---PLEFLFI 349
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  139 YNRKLNKSNIAFGGF-----EPSFDINMGOASALGF 170
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  350 FTAPLLIFYVHAAYSGFALLANFFNVITIGNSPSTGF 385
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
Q93RE6 PRELIMINARY; PRT; 427 AA.
AC Q93RE6;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Adhesin.
GN VAG.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RA Akeda Y., Honda T.;
RT "Vibrio parahaemolyticus adhesin gene."
  Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
  EMBL; AB047560; BAB59008.1; -.
  DR InterPro; IPR002528; MatE.
  DR Pfam; PF01554; UPF0013; 2.
  DR TIGRFAMs; TIGR00797; matE; 1.
  SQ SEQUENCE 427 AA; 45909 MW; 0B608B9DC70E3C71 CRC64;

Query Match 10.2%; Score 92; DB 2; Length 427;
Best Local Similarity 25.4%; Pred. No. 2.9;
Matches 50; Conservative 25; Mismatches 68; Indels 54; Gaps 10;

QY 5 LIVVLLAFCVSSFAQ-----ADDSKSAFNLAGKLLAYETSKDPIV----- 48
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  73 MVFAALAVCLFVFPPQIMGLTTDSQEVIELGSSYLIVISSASMFACVAVGLRAMH 132
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  49 -PFLNLFGLFGIGS--FAQGDILGGFLILGFDVAG-----ICLILTGAYLDIKALDKNA 100
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  133 QGLSTFFSGIGILSNVFLNWLIFGH--LGFPAUGITGAATATVISGA-IEVGCL---- 185
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  101 PKAAFKWTW-----GRGMMLAGAVTMVTRLEIILPPTFANSYNRKLNKSNIAFG 152
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  186 ----FCYLMKLNKHTAIFGWDIRASLVLDKIRFLSLSLPTTF-----NFLAWAGG 232
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 153 GFEPFSDINMGOASALG 169
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  233 LF--AYHAIMGAGVQG 247
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
Q8REW2 PRELIMINARY; PRT; 325 AA.
AC Q8REW2;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE CblB protein.
GN FN0975.
OS Fusobacterium nucleatum (subsp. nucleatum).

```



```
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=1886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusck G., Haselkorn R.,
RA Fonstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RL nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL: AE010604; AAL95171.1; -
KW Complete proteome.
SQ SEQUENCE 325 AA; 36409 MW; 9A53AE0C089DBAEA CRC64;

Query Match 10.0%; Score 90; DB 16; Length 325;
Best Local Similarity 23.2%; Pred. No. 3.2;
Matches 49; Conservative 29; Mismatches 85; Indels 48; Gaps 9;

4 FLIVLLAFVFFSFAQADSKSAF-----NLGAGEKLLAY----- 39
Db 76 YVIEIFFLFTLATKSLADEGKKVYRILKSGDIEKAKKELSYLVSRDNTLSLDKIIMSV 135
Qy 40 -ETSKKDPVFPFLNLFLGFGIGSFAOGDILGGFLILGFDAGVIGLILTGAYLDIKALD- 97
Db 136 VETAEINVGGFISPVFAF-VGSFFVVELFG-----KVVSALPFAFMYKAINTLDS 187
Qy 98 ----KNAPKAFAKFTWKGMMLAGVMTVTRTEIIPF--TFANSYNRKLNLSNIAFG 152
Db 188 MVGYKNEKYIDFGKVSARVDDIANFIPARLTGL--IFVPLSLILGYN--FKNSLKIFFR 243
Qy 153 GFEPSPDINMGQ-----ASALGPELSEKSY 178
Db 244 DRKKHSPNSQSESAYAGALGIGFGKISY 274

RESULT 12
Q9FDMO PRELIMINARY; PRT; 479 AA.
AC Q9FDMO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Metabolite transport protein.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
CX Zymomonas.
NCBI_TaxID=542;
RN SEQUENCE FROM N.A.
RC STRAIN=ZM4;
RA Lee H.J., Kang H.S.;
RT "Sequence analysis of fosmid clone 42D10 of Zymomonas mobilis ZM4.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AF212041; AAG02149.1; -
DR InterPro: IPR003663; CHO_transport.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF000083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane.
SQ SEQUENCE 479 AA; 52095 MW; 614B052B61AD2A12 CRC64;

Query Match 10.0%; Score 90; DB 2; Length 479;
Best Local Similarity 22.5%; Pred. No. 4.9;
Matches 29; Conservative 24; Mismatches 52; Indels 24; Gaps 4;
```

```
Qy 7 VVLLAFVFFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFLGFGIGSFAOG 66
Db 98 MIMIVTAIFGSVIAADAPTAFNLGFAVGGSSQIVPVY-----TAEIAPA 149
Qy 67 DILGGFLILGFDAGVIGLI---LFGAYLDIKALDKNAPKAFAKFTWKGMMLAGVMTAV 123
Db 150 DQRGRMVTFRYISIGLIGLAAGIVGAFILQ-----EETW-RTMFSVAIPA 196
Qy 124 TRLTEIIP 132
Db 197 LFCMMMLP 205

RESULT 13
Q97JE7 PRELIMINARY; PRT; 469 AA.
AC Q97JE7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Possible sugar-proton symporter.
GN CAC1339.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
CX Clostridiales; Clostridiaceae; Clostridium.
NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=2135325; PubMed=11466286;
RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RL bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007645; AAK79307.1; -
DR InterPro: IPR003663; CHO_transport.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRNSPORT.
DR TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 469 AA; 50482 MW; CDE33E2CD2A7B0DA CRC64;

Query Match 9.9%; Score 89.5; DB 16; Length 469;
Best Local Similarity 25.1%; Pred. No. 5.3;
Matches 44; Conservative 26; Mismatches 66; Indels 39; Gaps 8;

Qy 5 LIVVLLAFVFFSFAQADSKSAFNLGAGEKLLAYETSKKDPVFPFLNLFL- 58
Db 11 LLFTVLISCA-----AGIGLLYGYDTAVISGAIGFLKKLYNLSPAMOGF 55
Qy 59 GIGSFAOGDIIG---GFLIIGFDAGVIG--LILTGAYLDIKALDKNAPKAFAKFTWCK- 111
Db 56 VISSIMGGVILGVGFGSGLF---GDAIGRRKVLMLAALFAISAVISSISTSAFMLIFARI 112
Qy 112 --GMMLAGAVTMVTRTEIIPFTFANSYNRKLNLSNI-AFGGFEPSPDINMG 163
Db 113 VGGLIGMASALSVTYITECAPP-----SIRGLSSLYQLFTILGISTITFPFNIG 162

RESULT 14
Q9RTD0 PRELIMINARY; PRT; 312 AA.
AC Q9RTD0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide- transferase.
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 14:23:06 ; Search time 11 Seconds
(without alignments)
476.116 Million cell updates/sec

Title: US-09-508-487-21
Perfect score: 904
Sequence: 1 MNKFLIVLLAFCSFSAQ.....DINMQASALGFELSKKSY 178

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/FCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.5	9.0	858	US-09-255-829-22	Sequence 22, Appl
2	81.5	9.0	858	US-09-255-829-29	Sequence 29, Appl
3	81.5	9.0	1169	US-09-255-829-20	Sequence 20, Appl
4	72	8.0	344	US-09-134-001C-4175	Sequence 4175, Ap
5	72	8.0	451	US-08-679-635A-2	Sequence 2, Appl
6	72	8.0	451	US-09-419-163-2	Sequence 2, Appl
7	70.5	7.8	3491	US-07-642-734C-2	Sequence 2, Appl
8	70.5	7.8	3491	US-08-439-009A-2	Sequence 2, Appl
9	69.5	7.7	376	US-08-608-241-2	Sequence 2, Appl
10	69.5	7.7	376	US-08-922-182-2	Sequence 2, Appl
11	69.5	7.7	376	US-08-919-953-2	Sequence 2, Appl
12	69.5	7.7	376	US-09-192-983-2	Sequence 2, Appl
13	69.5	7.7	4928	US-09-036-987A-5	Sequence 5, Appl
14	69.5	7.7	4928	US-09-370-700-5	Sequence 5, Appl
15	69	7.6	931	US-08-624-655A-2	Sequence 2, Appl
16	69	7.6	2285	US-09-308-375-2	Sequence 2, Appl
17	68	7.5	280	US-08-855-140-1	Sequence 1, Appl
18	67.5	7.5	646	US-09-232-200-25	Sequence 25, Appl
19	67.5	7.5	646	US-09-232-200-32	Sequence 32, Appl
20	67.5	7.5	646	US-09-232-200-38	Sequence 38, Appl
21	67.5	7.5	646	US-09-232-200-43	Sequence 43, Appl
22	67.5	7.5	646	US-09-232-200-47	Sequence 47, Appl
23	67.5	7.5	646	US-09-232-200-59	Sequence 59, Appl
24	67.5	7.5	646	US-09-232-197-25	Sequence 25, Appl
25	67.5	7.5	646	US-09-232-197-32	Sequence 32, Appl
26	67.5	7.5	646	US-09-232-197-38	Sequence 38, Appl
27	67.5	7.5	646	US-09-232-197-43	Sequence 43, Appl

28	67.5	7.5	646	4	US-09-232-197-47	Sequence 47, Appl
29	67.5	7.5	646	4	US-09-232-197-59	Sequence 59, Appl
30	67.5	7.5	646	4	US-09-232-201-25	Sequence 25, Appl
31	67.5	7.5	646	4	US-09-232-201-32	Sequence 32, Appl
32	67.5	7.5	646	4	US-09-232-201-38	Sequence 38, Appl
33	67.5	7.5	646	4	US-09-232-201-43	Sequence 43, Appl
34	67.5	7.5	646	4	US-09-232-201-47	Sequence 47, Appl
35	67.5	7.5	646	4	US-09-232-201-59	Sequence 59, Appl
36	67.5	7.5	934	1	US-08-215-805A-80	Sequence 80, Appl
37	67.5	7.5	1584	4	US-09-251-645-6	Sequence 6, Appl
38	67	7.4	226	4	US-09-134-001C-5590	Sequence 5590, Ap
39	66.5	7.4	325	4	US-09-134-001C-3513	Sequence 3513, Ap
40	66.5	7.4	646	4	US-09-232-200-33	Sequence 33, Appl
41	66.5	7.4	646	4	US-09-232-200-65	Sequence 65, Appl
42	66.5	7.4	646	4	US-09-232-200-92	Sequence 92, Appl
43	66.5	7.4	646	4	US-09-232-197-33	Sequence 33, Appl
44	66.5	7.4	646	4	US-09-232-197-65	Sequence 65, Appl
45	66.5	7.4	646	4	US-09-232-197-92	Sequence 92, Ap

ALIGNMENTS

RESULT 1
US-09-255-829-22
; Sequence 22, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255.829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782.893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1581.0130002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 858 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-255-829-22

Query Match 9.0%; Score 81.5; DB 4; Length 858;
Best Local Similarity 23.8%; Pred. NO. 0.64;
Matches 36; Conservative 24; Mismatches 56; Indels 35; Gaps 6;


```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922,182
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/608,241
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 960296.93511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-922-182-2

Query Match 7.7%; Score 69.5; DB 2; Length 376;
Best Local Similarity 23.0%; Pred. No. 5.3;
Matches 43; Conservative 26; Mismatches 71; Indels 47; Gaps 7;

QY 23 DSKSAFNLGAGEKLLAY-----ETSKKDPVPFLLNLFLGFGIG----- 61
Db 120 DGTSRFSMLDGTPIHLHYMGCSFTSNVYVLPPIAVAKVRPDAPPDKICYIGCGVTTGIGAV 179
QY 62 -SFAOGDILGFLILGFDVAGI-----GLILTGAYLDKALDKNAPKAAFKWTWK----- 111
Db 180 INTAKVEIGAKAVYFGLGGIGLNVIOGLKLAGADMIIGVDLNNAK-----EWGERFGMT 234
QY 112 -----GMMLAGAYTMVAVTRTEI-IIPFTTFANSYNRK-LKNSLNIAFGGPEPSPDI 160
Db 235 HFVNPSEIDGDVAHLYNMKTPPDQIGGADYTFDCTGNVKNVQRQAELACHRGWQSQSIVI 294
QY 161 NMGOASA 167
Db 295 GVAPAGA 301

RESULT 11
US-08-919-953-2
; Sequence 2, Application US/08919953
; Patent No. 5837481
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy J
; APPLICANT: Barber, Robert D
; APPLICANT: Withuhn, Vernon
; TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
; TITLE OF INVENTION: SENSING AND REMEDIATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Plinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,953
; FILING DATE:
; CLASSIFICATION: 435

```

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/608,241
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Seay, Nicholas J
;; REGISTRATION NUMBER: 27,386
;; REFERENCE/DOCKET NUMBER: 960296, 93511
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 608-251-5000
;; TELEFAX: 608-251-9166
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 376 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-919-953-2

Query Match 7.7%; Score 69.5; DB 2; Length 376;
Best Local Similarity 23.0%; Pred. No. 5.3;
Matches 43; Conservative 26; Mismatches 71; Indels 47; Gaps 7;

QY 23 DSKSAFNLGAGEKLLAY-----ETSKKDPVLPFLNLFGLFGIG----- 61
DB 120 DTSRFSMLDGPILHYMGCSFNSVTVLPETAVAKVRDPADFDKICYIGCGVTTGIGAV 179
QY 62 -SFAQGDILGGLILGFDVAVGI-----GLILTGAYLIDIKALDNAPKAAFKWTGK----- 111
DB 180 INTAKVEICAKAVVFLGIGLNVLIQGLKLAGADMIIIGVDLNNNAK-----EMGERFGMT 234
QY 112 -----GMMLAGAVTMAVTRLEI-IIPPTFANSYNRK-LKNSLNIAFGGFEPSPDI 160
DB 235 HFVNPSEIDGDVVVAHLVNNMTKTPFDQIGADYTFDCTGNVKVMRQALEACHRGWGQSIVI 294
QY 161 NMQQASA 167
DB 295 GVAPAGA 301

RESULT 12
US-09-192-983-2
; Sequence 2, Application US/09192983A
; Patent No. 6242244
; GENERAL INFORMATION:
; APPLICANT: Donohue, Timothy
; APPLICANT: Barber, Robert
; APPLICANT: Witthuhn, Vernon
; TITLE OF INVENTION: Microbial System for Formaldehyde Sensing and
; TITLE OF INVENTION: Remediation
; FILE REFERENCE: 960296.95505
; CURRENT APPLICATION NUMBER: US/09/192,983A
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/919,953
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 08/608,241
; EARLIER FILING DATE: 1996-02-28
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Rhodobacter sphaeroides
US-09-192-983-2

Query Match 7.7%; Score 69.5; DB 4; Length 376;
Best Local Similarity 23.0%; Pred. No. 5.3;
Matches 43; Conservative 26; Mismatches 71; Indels 47; Gaps 7;

QY 23 DSKSAFNLGAGEKLLAY-----ETSKKDPVLPFLNLFGLFGIG----- 61
DB 120 DTSRFSMLDGPILHYMGCSFNSVTVLPETAVAKVRDPADFDKICYIGCGVTTGIGAV 179
QY 62 -SFAQGDILGGLILGFDVAVGI-----GLILTGAYLIDIKALDNAPKAAFKWTGK----- 111
```

```
DB 180 INTAKVEICAKAVVFLGIGLNVLIQGLKLAGADMIIIGVDLNNNAK-----EMGERFGMT 234
QY 112 -----GMMLAGAVTMAVTRLEI-IIPPTFANSYNRK-LKNSLNIAFGGFEPSPDI 160
DB 235 HFVNPSEIDGDVVVAHLVNNMTKTPFDQIGADYTFDCTGNVKVMRQALEACHRGWGQSIVI 294
QY 161 NMQQASA 167
DB 295 GVAPAGA 301

RESULT 13
US-09-036-987A-5
; Sequence 5, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4928 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-987A-5

Query Match 7.7%; Score 69.5; DB 4; Length 4928;
Best Local Similarity 25.2%; Pred. No. 1.9e+02;
Matches 39; Conservative 25; Mismatches 48; Indels 43; Gaps 8;

QY 46 PIVPFLNLFGLFGIGS-----FAQGDILGGLILGFDFA-----VGIGLILTGAYLIDIKAL 96
DB 2231 PVPVPMVS-----GKTPEALSQAQADALMSYLSNRPDASPRDIGYSLAVTRPALDHRV 2283
QY 97 DKNAPKAAFKWTGKMMIAGAVTMAVTR-LFEII-----IIPPTFANSYNRK--KLKN 145
DB 2284 VLGDARAA-----LLPGLKALAVSNDAAAEVITGTAAAGPVGVFSGQGWPGMGS 2334
QY 146 SUNIAFGGFEPSPF-----DINMGQASALGFELS 173
```


Db 2335 GLHSAPFVFADAFDEACCELDHAHLGQMARLRDVL 2369

RESULT 14

US-09-370-700-5
; Sequence 5, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes for Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4928
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-5

Query Match 7.7%; Score 69.5; DB 4; Length 4928;
Best Local Similarity 25.2%; Pred. No. 1.9e+02;
Matches 39; Conservative 25; Mismatches 48; Indels 43; Gaps 8;

Qy 46 PIVPELNLNLFGLFGIGS-----FAGDILGGFLILGFDA---VIGLILTGAYLDIKAL 96
Db 2231 PVVPMVMS-----GKTPEALSQAQADALMSYLSNRVDASPRDIGYSLAVTRPALDHRV 2283
Qy 97 DKNAPKAAFKWTGKGMLAGAVTMAVTR-LTEII-----IPFTFANSYNR--KLKN 145
Db 2284 VLGDRAA-----LLPLKALAVSNDAAEVITGTRAGPVGVFVSGGGQWPGMG 2334
Qy 146 SLNIAFGGFEPF-----DINMQASALGFELS 173
Db 2335 GLHSAPFVFADAFDEACCELDHAHLGQMARLRDVL 2369

RESULT 15

US-08-624-655A-2
; Sequence 2, Application US/08624655A
; Patent No. 6323005
; GENERAL INFORMATION:
; APPLICANT: DABAN, MONTERRAT
; APPLICANT: MEDRANO, ANDRES
; APPLICANT: ESPUNA, ENRIC
; APPLICANT: QUEROL, ENRIQUE
; TITLE OF INVENTION: TRANSFERRIN-BINDING PROTEIN 1 (Tbpl) GENE OF
; TITLE OF INVENTION: Actinobacillus pleuropneumoniae, ITS USE TO PREPARE
; TITLE OF INVENTION: PRODUCTS FOR THE UTILIZATION IN VACCINES FOR
; TITLE OF INVENTION: PLEUROPNEUMONIA AND AS DIAGNOSTIC REAGENTS
; FILE REFERENCE: P00740
; CURRENT APPLICATION NUMBER: US/08/624,655A
; CURRENT FILING DATE: 1996-03-22
; PRIOR APPLICATION NUMBER: 95 00592
; PRIOR FILING DATE: 1995-03-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Actinobacillus pleuropneumoniae
US-08-624-655A-2

Query Match 7.6%; Score 69; DB 4; Length 931;

Best Local Similarity 21.2%; Pred. No. 21;
Matches 44; Conservative 27; Mismatches 69; Indels 68; Gaps 9;
Qy 9 LLAFCVFSSFAQADDSSKSAFNL-----GAGEKLLAYETSKKOPIV 48
Db 7 LISLALLSLFAVQSYAQAVOLNDVYVTGTTKKKAHKKENEVTGLGKVVKTPDTLSKEOV- 65
Qy 49 PFLNLFLGF-----GIGSFAQG-DILGGFLILGFDAVIGLILTG----- 88
Db 66 -----LGIRDLTRYDPGISVVEQGRGATTGYSIRGVDRNRVGLALDGLPQIQSYVSQY 118
Qy 89 -----AYLDIKALDKNAPKAAFKWTGKGMLAGAVTMAVTRLTETIIP----- 132
Db 119 SRSSSGAINETEYENLRSI--QISKGASSSEFGSG-SLGGSVQFRTKEVSDIIRPGQSWG 175
Qy 133 FTFANSYNRKLKNSLN-IAFCGFEPSPD 159
Db 176 LDTKSAYSSKNQOOWLNSLAFAGTHNGFD 203
Search completed: March 14, 2003, 14:27:47
Job time : 16 secs

THIS PAGE BLANK (USPTO)